

307.801

A MAGYAR TUDOMÁNYOS AKADÉMIA

MATEMATIKAI
KUTATÓ INTÉZETÉNEK
KÖZLEMÉNYEI

IV. ÉVFOLYAM, 1. FÜZET

1959

★

ТРУДЫ
МАТЕМАТИЧЕСКОГО
ИНСТИТУТА
АКАДЕМИИ НАУК ВЕНГРИИ
ТОМ IV., ВЫПУСК 1.
1959

★

PUBLICATIONS
OF THE
MATHEMATICAL INSTITUTE
OF THE
HUNGARIAN ACADEMY OF SCIENCES
VOLUME IV., FASC. 1.
1959



1959

2

TARTALOMJEGYZÉK

SCHNELL E.: Egy feltételes határelosztástételről	3
BRÓDY A.: Vállalati és népgazdasági megtakarítás	11
EGERVÁRY J.: Kombinatorikus módszer a szállítási probléma megoldására ...	15
VINCZE I.: Kétdimenziós eloszlás és határelosztástételek a rendezett minták elméletében, II.	29
ERDŐS P.—RÉNYI A.: Véges sokaságból vett mintákra vonatkozó centrális határelosztástétel	49
DÉNES J.: Permutációk minimális számú transzpozíció szorzatára bontásainak száma és ezek kapcsolata a gráfelmélettel	63
RÉNYI A.: Megjegyzések a „fák” elméletéhez	73
HANDSCOMB, D. C.: Rényi Alfréd egy problémájáról	87
SZ.-NAGY B.: Teljesen folytonos operátorokról, amelyek iteráltjai egyenletesen korlátosak	89
GEHÉR L.: A véges dimenziójú terek hasonló lineáris transzformációiról.....	95

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI
KUTATÓ INTÉZETÉNEK
KÖZLEMÉNYEI

IV. ÉVFOLYAM

1959

★

ТРУДЫ
МАТЕМАТИЧЕСКОГО
ИНСТИТУТА
АКАДЕМИИ НАУК ВЕНГРИИ

ТОМ IV.

1959

★

PUBLICATIONS
OF THE
MATHEMATICAL INSTITUTE
OF THE
HUNGARIAN ACADEMY OF SCIENCES
VOLUME IV.
1959



1959

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI KUTATÓ INTÉZETÉNEK
KÖZLEMÉNYEI

„A MAGYAR TUDOMÁNYOS AKADÉMIA ALKALMAZOTT MATEMATIKAI INTÉZETÉNEK KÖZLEMÉNYEI”

című kiadványsorozat folytatása

SZERKESZTI: RÉNYI ALFRÉD

TECHNIKAI SZERKESZTŐK: RÉVÉSZ PÁL, BOGNÁR KATALIN

A SZERKESZTŐSÉG CÍME: MATEMATIKAI KUTATÓ INTÉZET, BUDAPEST, V. REÁLTANODA U. 13/15.

A MATEMATIKAI KUTATÓ INTÉZET KÖZLEMÉNYEI az Intézet tudományos eredményeit tartalmazó és egyéb matematikai, valamint a matematika gyakorlati alkalmazásával kapcsolatos dolgozatokat közölnek. A folyóirat minden évfolyama 4 füzetből áll és körülbelül 30 nyomdai terjedelmű. A dolgozatok vagy valamelyik világnyelven jelennek meg, magyar és még egy világnyelven írt részletes kivonattal, vagy pedig magyarul, két világnyelven írt részletes kivonattal. Közlésre szánt dolgozatokat kérjük 2 gépelt példányban kivonattal ellátva a szerkesztő címére küldeni (Budapest V. Reáltanoda u. 13/15.).

A folyóirat, amelynek első évfolyama 1956-ban jelent meg, folytatása a „Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei” című kiadványsorozatnak, amelynek összesen három kötete jelent meg: I. kötet (1952), II. kötet (1953), III. kötet (1954).

A MATEMATIKAI KUTATÓ INTÉZET KÖZLEMÉNYEINEK előfizetési ára évfolyamonként belföldi címre 50,— Ft, külföldi címre 70,— Ft. (Kötetenként 5 \$) Belföldi megrendelések az Akadémiai Kiadón keresztül adhatók fel (Budapest V. Alkotmány u. 21., Magyar Nemzeti Bank egyszámlaszám: 15.915.111—46), külföldi megrendelések a Posta Központi Hírlap Iroda útján eszközölhetők (Budapest V. József nádor tér 1., Magyar Nemzeti Bank egyszámlaszám: 61.257). A folyóirat egyes füzetei 15.— Ft-os árban az Akadémiai Könyvesboltban kaphatók (Budapest V. Váci u. 22.) Cserkeapcsolatok felvétele érdekében kérjük az Intézet Könyvtárához fordulni (Budapest V. Reáltanoda u. 13/15.)

ТРУДЫ
МАТЕМАТИЧЕСКОГО ИНСТИТУТА

[АКАДЕМИИ НАУК ВЕНГРИИ]

[Продолжение издания]

„A MAGYAR TUDOMÁNYOS AKADÉMIA ALKALMAZOTT MATEMATIKAI INTÉZETÉNEK KÖZLEMÉNYEI”

РЕДАКТОР: ALFRÉD RÉNYI

ТЕХНИЧЕСКИЕ РЕДАКТОРЫ: ПÁL RÉVÉSZ, KATALIN BOGNÁR

АДРЕС: МАТЕМАТИЧЕСКИЙ ИНСТИТУТ, БУДАПЕСТ, V. РЕАЛТАНОДА U. 13/15., ВЕНГРИЯ

В ТРУДАХ МАТЕМАТИЧЕСКОГО ИНСТИТУТА печатаются статьи, содержащие результаты научно-исследовательской работы Института, и другие математические работы, а также статьи, связанные с практическими приложениями математики. Каждый том журнала выходит в 4 выпусках и содержит приблизительно 30 печатных листов. Статьи публикуются либо на каком-нибудь мировом языке с подробным резюме на венгерском и каком-нибудь другом мировом языке, либо на венгерском языке с подробным резюме на двух мировых языках. Работы, предназначенные для опубликования, просим посылать в двух напечатанных на машинке экземплярах вместе с резюме в адрес редакции (Будапест, V. Реáltanoda u. 13/15.).

Журнал, первый том которого вышел в 1956 году, является продолжением издания «A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei» (Труды Института Прикладной Математики Академии Наук Венгрии) вышедшего в трёх томах: Том I. (1952), Том II. (1953), Том III. (1954).

Цена подписки на ТРУДЫ МАТЕМАТИЧЕСКОГО ИНСТИТУТА в Венгрии 50 форинтов, на заграничный адрес 70 форинтов (\$5 — за каждый том). Местные заказы принимает Издательство Академии Наук (Budapest V. Alkotmány u. 21., счёт Венгерского Национального Банка 15.915.111—46), заграничные заказы принимает Центральная Журнальная Контора Почты (Budapest V. József nádor tér 1., счёт Венгерского Национального Банка 61.257). По поводу отношения обмена просим обращаться к Библиотеке Института (Budapest V. Reáltanoda u. 13/15. Венгрия).

PUBLICATIONS
OF THE MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES

continuing the series

„A MAGYAR TUDOMÁNYOS AKADÉMIA ALKALMAZOTT MATEMATIKAI INTÉZETÉNEK KÖZLEMÉNYEI”

EDITOR: ALFRÉD RÉNYI

TECHNICAL EDITORS: PÁL RÉVÉSZ, KATALIN BOGNÁR

ADDRESS: MATHEMATICAL INSTITUTE, BUDAPEST, V., REÁLTANODA U. 13/15, HUNGARY

The PUBLICATIONS OF THE MATHEMATICAL INSTITUTE are publishing papers containing the results of scientific work of the Institute and other mathematical papers and papers on the practical applications of mathematics. The journal is published quarterly, 4 fasciuli are forming a volume consisting of about 30 printed lists. The papers appear either in a world-language with abstract in Hungarian and in an other world-language, or in Hungarian with abstracts in two world-languages. Papers intended for publication in the journal should be sent to the editor in 2 typewritten copies, with an abstract.

The journal, the first-volume of which appeared in 1956, continues the series „A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei” (Publications de l'Institut des Mathématiques Appliquées de l'Académie des Sciences de Hongrie) of which 3 volumes were published altogether: Volume I. (1952), Volume II. (1953), Volume III. (1954).

The price of a volume of the PUBLICATIONS OF THE MATHEMATICAL INSTITUTE is 50,— Ft to an address in Hungary and 70,— Ft (\$ 5.— The Volume) to abroad. Subscriptions can be made at the Académie Publishing House in Hungary (Budapest V. Alkotmány u. 21., single account number in the Hungarian National Bank 15.915.111—46) resp. at the Posta Központi Hírlap Iroda from abroad (Budapest V. József nádor tér 1., single-account number in the Hungarian National Bank 61.257). For establishing exchange relations please write to the Library of the Mathematical Institute (Budapest V. Reáltanoda u. 13/15., Hungary).

TARTALOMJEGYZÉK

ÁDÁM A.: Kétpólusú elektromos hálózatokról, IV.	183
ADLER GY.: A hővezetés peremértékfeladatainak egy új típusa	109
ADLER GY.: Szemcsés anyagok ellenáramú hűtése	327
ADLER GY.—FREUD G.: A Mikusióski-féle operátorszámítás alkalmazása egy parciális differenciálegyenlet megoldására	367
ALMÁR I.—BALÁZS B.: Közelítő módszer a Hold mellett elhaladó úrrakéta pályájának megszerkesztésére	129
BÉKÉSSY A.—TÓTH K.: Megjegyzés Almár I. és Balázs B.: „Közelítő módszer a Hold mellett elhaladó úrrakéta pályájának megszerkesztésére” c. cikkéhez	143
BOSZNYAI Á.: Többszabadságfokú lengőrendszer gerjesztett lengései vizsgálatának egy módszere s néhány alkalmazása	299
BRÓDY A.: Válatati és népgazdasági megtakarítás	11
CSÁKI E.: A Wilcoxon-statisztika két módosításáról	313
CSORDÁS L.—MENCZEL GY.: Megjegyzések Medgyessy Pál: „Egy mechanikai függvény szintetizátor” című cikkéhez	149
DÉNES J.: Permutációk minimális számú transzpozíció szorzatára bontásainak száma és ezek kapcsolata a gráfelmélettel.	63
EGERVÁRY J.: Kombinatorikus módszer a szállítási probléma megoldására.	15
EGERVÁRY J. †	107
ERDŐS P.—RÉNYI A.: Véges sokaságból vett mintákra vonatkozó centrális határeloszlástétel	49
FÉNYES T.: A Mikusióski-féle operátorszámítás alkalmazása állandó együththatós retardált differenciálegyenletek megoldására	191
FINNEY, D. J.: A kísérletezés rentabilitása	203
GEHÉR L.: A véges dimenziójú terek hasonló lineáris transzformációiról.	95
GRÄTZER GY.—SCHMIDT E. T.: Egy asszociativitási tétel alternatív gyűrűkre.	259
HANDSCOMB, D. C.: Rényi Alfréd egy problémájáról	87
HASZPRA O.: Néhány megjegyzés nomogramok leolvasási hibáihoz	321
KOSIK P.—SALLAY M.—ZIMÁNYI M.: Hővezetési problémák összetett peremfeltételek esetén	377
KREKÓ B.: Lineáris egyenletrendszerek megoldása szimplex módszerrel.	265
LINNİK, JU. V.: Öt előadás a valószínűségszámítás és a számelmélet bizonyos problémáiról	225
RÉNYI A.: Megjegyzések a „fák” elméletéhez	73
RÉNYI A.: Autoklávok soros és párhuzamos kapcsolásáról és a keverés elméletéről	155
RÉNYI A.: Összefüggő gráfokról, I.	385
RÉNYI A.: Szummációs eljárások valószínűségszámítási interpretációja.	389
RÓZSA P.—SÁRKÁNY GY.: Ellenáramú szétválasztó vegyipari alapl műveletek elméleti fokozatszámának meghatározásáról, II.	277
SCHNELL E.: Egy feltételes határeloszlástételről	3
SZ.-NAGY B.: Teljesen folytonos operátorokról, amelyek iteráltjai egyenletesen korlátosak	89
SZÜSZ P.: Feltételeken konvergencia sorok átrendezéséről.	167
VAS É.: A szekvenciális valószínűségi hánnyados próba efficienciájáról.	171
VINCZE I.: Kétdimenziós eloszlás és határeloszlástételek a rendezett minták elméletében, II.	29

RÖVID KÖZLEMÉNYEK

ADLER GY.: Egy megjegyzés az úrhajózásról	197
--	-----

СОДЕРЖАНИЕ

ÁDÁM, A.: О двухполюсных электрических сетях, IV.	183
ADLER, GY.: Один новый тип граничных задач теплопроводности	109
ADLER, GY.: Охлаждение зернистых веществ встречным потоком	327
ADLER, GY. — FREUD, G.: Применение операционного исчисления от Mikusiński к решению одного уравнения в частных производных	367
ALMÁR, I. — BALÁZS, B.: Приближенный метод расчета траектории космической ракеты проходящей вблизи Луны	129
BÉKÉSSY, A. — TÓTH, K.: Замечание к статье I. Almár и B. Balázs: „Приближённый метод расчета траектории космической ракеты проходящей вблизи Луны”	143
BOSZNAI, A.: Один метод исследования возбужденных колебаний колеблющихся систем с несколькими степенями свободы и несколько его приложений	299
BRÓDY, A.: Об экономии предприятия и народного хозяйства	11
CSÁKI, E.: О двух видоизменениях статистики Wilcoxon-a	313
CSORDÁS, L. — MENCSEI, GY.: Замечания к статье P. Medgyessy: „Один механиче- ский функциональный синтезатор”	149
DÉNES, J.: Число разложений перестановок в произведение наименьшего числа транспозиций и их связь с теорией графов	63
EGERVÁRY, J.: Комбинаторный метод для решения проблемы перевозки	15
EGERVÁRY, J. †	107
ERDŐS, P. — RÉNYI, A.: Центральная предельная теорема для выборках взятых из конечных множеств	49
FÉNYES, T.: Применение операционного исчисления к решению обыкновенных раз- ностно-дифференциальных уравнений	191
FINNEY, D. J.: О рентабельности экспериментирования	203
GENÉR, L.: О подобных линейных преобразованиях конечномерных пространств	95
GRÄTZER, GY. — SCHMIDT, E. T.: Одна теорема об ассоциативности для альтерна- тивных колец	259
HANDSCOMB, D. C.: Об одной проблеме A. Rényi	87
HASZPRA, O.: Несколько замечаний в связи с погрешностями номограмм	321
KOSIK, P. — SALLAY, M. — ZIMÁNYI, M.: Проблемы теплопроводности в случае сложных условий	377
КРЕКÓ, В.: Решение систем линейных уравнений методом симплекса	265
Линник, Ю. В.: Пять лекций о некоторых вопросах теории чисел и теории веро- ятностей	225
RÉNYI, A.: Замечания к теории «деревьев»	73
RÉNYI, A.: О последовательном и параллельном сцеплении автоклавов и о теории смешения	155
RÉNYI, A.: О связанных графах, I.	385
RÉNYI, A.: Методы суммирования и их теоретико-вероятностная интерпретация	389
RÓZSA, P. — SÁRKÁNY, GY.: Определение числа теоретических ступеней противоточных разделяющих процессов химической технологии, II.	277
SCHNELL, E.: Об одной теореме об условном предельном распределении для оди- накового распределенных независимых случайных величин	3
Sz.-NAGY, B.: О вполне непрерывных операторах с равномерно ограниченными итерированными	89
Szűsz, P.: О перестановке членов условно сходящихся рядов	167
VAS, É.: Об эффективности секвенциального критерия отношения вероятностей	171
VINCZE, I.: Двумерные теоремы о распределении и предельном распределении в теории вариационных рядов, II.	29

КРАТКИЕ СООБЩЕНИЯ


ADLER, GY.: Замечание о межпланетных путешествиях	197
---	-----

INDEX

ADÁM, A.: Über zweipolige elektrische Netze, IV.....	183
ADLER, G.: Un type nouveau des problèmes aux limites de la conduction de la chaleur	109
ADLER, G.: Réfrigération de retour des matières granuleuses.....	327
ADLER, G.—FREUD, G.: Una applicazione del calcolo degli operatori di Mikusiński per la risoluzione d'una equazione alle derivate parziali.....	367
ALMÁR, I.—BALÁZS, B.: Approximate method of plotting the orbit of a space rocket passing near the Moon	129
BÉKÉSSY, A.—TÓTH, K.: Remark on the paper of I. Almár and B. Balázs: „Aproxi- mate method of plotting the orbit of a space rocket passing near the Moon”	143
BOSZNAVY, Á.: Eine Methode zur Untersuchung der erzwungenen Schwingungen von Schwingungssystemen mit mehreren Freiheitsgraden und deren Anwendungen	299
BRÓDY, A.: Savings, computed on a price basis and savings computed on the basis of a technological matrix	11
CSÁKI, E.: On two modifications of the Wilcoxon's test.....	313
CSORDÁS, L.—MENCZEL, G.: Remarks on the paper of P. Medgyessy: „A mechanical functional synthesizer”	149
DÉNES, J.: The representation of a permutation as the product of a minimal number of transpositions, and its connection with the theory of graphs.....	63
EGERVÁRY, J.: Combinatorial method for solving the transport problem.....	15
EGERVÁRY, J. †	107
ERDŐS, P.—RÉNYI, A.: On the central limit theorem for samples from a finite population	49
FÉNYES, T.: Anwendung der Operatorenrechnung von Mikusiński zur Lösung gewöhnlicher Differentialdifferenzengleichungen	191
FINNEY, D. J.: The economic efficiency of experimentation	203
GEHÉR, L.: Über ähnliche lineare Transformationen in endlichdimensionalen Räu- men	95
GRÁTZER, GY.—SCHMIDT, E. T.: An associativity theorem for alternative rings....	259
HANDSCOMB, D. C.: A note on a problem of Rényi	87
HASZPRA, O.: Einige Bemerkungen zur Berechnung der Nomogrammbaufehler..	321
KOSIK, P.—SALLAY, M.—ZIMÁNYI, M.: Problèmes de la propagation de la chaleur avec des conditions aux limites composées	377
KREKÓ, B.: Solution of a system of linear equations with the simplex method	265
LINNIK, JU. V.: Five lectures on some topics in number theory and probability theory	225
RÉNYI, A.: Some remarks on the theory of trees	73
RÉNYI, A.: On serial and parallel coupling of autoclaves and on the theory of mixing	155
RÉNYI, A.: On connected graphs, I.	385
RÉNYI, A.: Summation methods and probability theory	389
RÓZSA, P.—SÁRKÁNY, GY.: Determination of the theoretical stage-number of coun- tercurrent distributive chemical unit operations, II.	277
SCHNELL, E.: On a conditional limiting distribution theorem.....	3
SZ.-NAGY, B.: Completely continuous operators with uniformly bounded iterates	89
SZÜSZ, P.: Über die Umordnung bedingt konvergenter Reihen.....	167
VAS, É.: On the efficiency of the sequential probability ratio test.....	171
VINCZE, I.: On some distributions and joint limiting distributions in the theory of order statistics, II.	29

SHORT COMMUNICATIONS

ADLER, GY.: Une remarque sur la communication intercosmique.....	197
--	-----



A kiadásért felel az Akadémiai Kiadó igazgatója

50.431/60 — Akadémiai Nyomda, Budapest — Felelős vezető: Bernát György

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI
KUTATÓ INTÉZETÉNEK
KÖZLEMÉNYEI

IV. ÉVFOLYAM, 1. FÜZET

1959

★

ТРУДЫ
МАТЕМАТИЧЕСКОГО
ИНСТИТУТА
АКАДЕМИИ НАУК ВЕНГРИИ
ТОМ IV., ВЫПУСК 1.
1959

★

PUBLICATIONS
OF THE
MATHEMATICAL INSTITUTE
OF THE
HUNGARIAN ACADEMY OF SCIENCES
VOLUME IV., FASC. 1.
1959



1959

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI KUTATÓ INTÉZETÉNEK
KÖZLEMÉNYEI

„A MAGYAR TUDOMÁNYOS AKADÉMIA ALKALMAZOTT MATEMATIKAI INTÉZETÉNEK KÖZLEMÉNYEI”
című kiadványsorozat folytatása

SZERKESZTI: RÉNYI ALFRÉD

TECHNIKAI SZERKESZTŐK: RÉVÉSZ PÁL, WEISZ KATALIN

A SZERKESZTŐSÉG CÍME: MATEMATIKAI KUTATÓ INTÉZET, BUDAPEST, V. REÁLTANODA U. 13/15.

A MATEMATIKAI KUTATÓ INTÉZET KÖZLEMÉNYEI az Intézet tudományos eredményeit tartalmazó és egyéb matematikai, valamint a matematika gyakorlati alkalmazásával kapcsolatos dolgozatokat közölnek. A folyóirat minden évfolyama 4 füzetből áll és körülbelül 30 nyomdai terjedelmű. A dolgozatok vagy valamelyik világnyelven jelennek meg, magyar és még egy világnyelven írt részletes kivonattal, vagy pedig magyarul, két világnyelven írt részletes kivonattal. Közlésre szánt dolgozatokat kérjük 2 gépelt példányban kivonattal ellátva a szerkesztő címére küldeni (Budapest V. Reáltanoda u. 13/15.).

A folyóirat, amelynek első évfolyama 1956-ban jelent meg, folytatása a „Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei” című kiadványsorozatnak, amelynek összesen három kötete jelent meg: I. kötet (1952), II. kötet (1953), III. kötet (1954).

A MATEMATIKAI KUTATÓ INTÉZET KÖZLEMÉNYEINEK előfizetési ára évfolyamonként belföldi címre 50,— Ft, külföldi címre 70,— Ft. (Kötetenként 5 \$) Belföldi megrendelések az Akadémiai Kiadón keresztül adhatók fel (Budapest V. Alkotmány u. 21., Magyar Nemzeti Bank egyszámlaszám: 15.915.111—46), külföldi megrendelések a Posta Központi Hírlap Iroda útján eszközölhetők (Budapest V. József nádor tér 1., Magyar Nemzeti Bank egyszámlaszám: 61.257). A folyóirat egyes füzetei 15,— Ft-os árban az Akadémiai Könyvesboltban kaphatók (Budapest V. Váci u. 22.) Cserekapcsolatok felvétele érdekében kérjük az Intézet Könyvtárához fordulni (Budapest V. Reáltanoda u. 13/15.)

ТРУДЫ
МАТЕМАТИЧЕСКОГО ИНСТИТУТА
АКАДЕМИИ НАУК ВЕНГРИИ

Продолжение издания

„A MAGYAR TUDOMÁNYOS AKADÉMIA ALKALMAZOTT MATEMATIKAI INTÉZETÉNEK KÖZLEMÉNYEI”

РЕДАКТОР: ALFRÉD RÉNYI

ТЕХНИЧЕСКИЕ РЕДАКТОРЫ: ПÁL RÉVÉSZ, KATALIN WEISZ

АДРЕС: МАТЕМАТИЧЕСКИЙ ИНСТИТУТ, BUDAPEST, V. REÁLTANODA U. 13/15., ВЕНГРИЯ

В ТРУДАХ МАТЕМАТИЧЕСКОГО ИНСТИТУТА печатаются статьи, содержащие результаты научно-исследовательской работы Института, и другие математические работы, а также статьи, связанные с практическими приложениями математики. Каждый том журнала выходит в 4 выпусках и содержит приблизительно 30 печатных листов. Статьи публикуются либо на каком-нибудь мировом языке с подробным резюме на венгерском и каком-нибудь другом мировом языке, либо на венгерском языке с подробным резюме на двух мировых языках. Работы, предназначенные для опубликования, просим посылать в двух напечатанных на машинке экземплярах вместе с резюме в адрес редакции (Budapest, V. Reáltanoda u. 13/15.).

Журнал, первый том которого вышел в 1956 году, является продолжением издания «A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei» (Труды Института Прикладной Математики Академии Наук Венгрии) вышедшего в трёх томах: Том I. (1952), Том II. (1953), Том III. (1954).

Цена подписки на ТРУДЫ МАТЕМАТИЧЕСКОГО ИНСТИТУТА в Венгрии 50 форинтов, на заграничный адрес 70 форинтов (\$ 5.—) за каждый том. Местные заказы принимает Издательство Академии Наук (Budapest V. Alkotmány u. 21., счёт Венгерского Национального Банка 15.915.111—46), заграничные заказы принимает Журнальная Контора Почты (Budapest V. József nádor tér 1., счёт Венгерского Национального Банка 61.257). По поводу отношения обмена просим обращаться к Библиотеке Института (Budapest V. Reáltanoda u. 13/15., Венгрия).

PUBLICATIONS
OF THE MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES

continuing the series

„A MAGYAR TUDOMÁNYOS AKADÉMIA ALKALMAZOTT MATEMATIKAI INTÉZETÉNEK KÖZLEMÉNYEI”

EDITOR: ALFRÉD RÉNYI

TECHNICAL EDITORS: PÁL RÉVÉSZ, KATALIN WEISZ

ADDRESS: MATHEMATICAL INSTITUTE, BUDAPEST, V. REÁLTANODA U. 13/15, HUNGARY

The PUBLICATIONS OF THE MATHEMATICAL INSTITUTE are publishing papers containing the results of scientific work of the Institute and other mathematical papers and papers on the practical applications of mathematics. The journal is published quarterly, 4 fasciculi are forming a volume consisting of about 30 printed lists. The papers appear either in a world-language with abstract in Hungarian and in an other world-language, or in Hungarian with abstracts in two world-languages. Papers intended for publication in the journal should be sent to the editor in 2 typewritten copies, with an abstract.

The journal, the first-volume of which appeared in 1956, continues the series «A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei» (Publications of l'Institut des Mathématiques Appliquées de l'Académie des Sciences de Hongrie) of which 3 volumes were published altogether: Volume I. (1952), Volume II. (1953), Volume III. (1954).

The price of a volume of the PUBLICATIONS OF THE MATHEMATICAL INSTITUTE is 50,— Ft to an address in Hungary and 70,— Ft (\$ 5.— The Volume) to abroad. Subscriptions can be made at the Academic Publishing House in Hungary (Budapest V. Alkotmány u. 21., single account number in the Hungarian National Bank 15.915.111—46) resp. at the Posta Központi Hírlap Iroda from abroad (Budapest V. József nádor tér 1., single-account number in the Hungarian National Bank 61.257). For establishing exchange relations please write to the Library of the Mathematical Institute (Budapest V. Reáltanoda u. 13/15., Hungary).

ON A CONDITIONAL LIMITING DISTRIBUTION THEOREM

by

EDIT SCHNELL

This paper deals with the following problem: under what conditions will the conditional limiting distribution of the non-standardized sum of independent and identically distributed random variables be uniform on the whole real axis (see RÉNYI [4]). By other words under what conditions will the relation

$$(1) \quad \lim_{n \rightarrow \infty} \mathbf{P} \left\{ a_1 \leq \sum_{i=1}^n \xi_i < b_1 \mid a_2 \leq \sum_{i=1}^n \xi_i < b_2 \right\} = \frac{b_1 - a_1}{b_2 - a_2} \quad (a_2 \leq a_1 < b_1 < b_2)$$

be valid, where $\xi_1, \xi_2, \dots, \xi_n, \dots$ are independent and identically distributed random variables.

In their paper [3] C. KALLIANPUR and H. ROBBINS have given a sufficient condition for the validity of (1) which is scarcely more general than the assumption that the ξ_i -s are nonlattice distributed (see [2], p. 58.), they have 0 expectation and finite variance¹ and:

$$(2) \quad |\mathbf{M}\{\xi_i^3\}| < \infty \quad i = 1, 2, \dots$$

Under such conditions the validity of (1) may be derived from a well-known estimation of ESSEEN (see [2], p. 213). According to this estimation under the above conditions

$$(3) \quad F_n(x) = \Phi(x) + \frac{C}{\sqrt{n}} \{ \Phi'(x) (1 - x^2) + o(1) \}$$

uniformly in x , where $F_n(x)$ is the distribution function of the random variable

$$(4) \quad \zeta_n = \frac{\eta_n}{\sqrt{n}} = \frac{\sum_{i=1}^n \xi_i}{\sqrt{n}},$$

$$(5) \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$$

¹ We may suppose that the variance of the variables is 1. If in this case the theorem applies to some pair of intervals $[a_1, b_1], [a_2, b_2]$, so it applies to the pair $[\sigma a_1, \sigma b_1], [\sigma a_2, \sigma b_2]$ for the variables: $\xi_i = \sigma \xi_i$.

and C is the constant

$$(6) \quad C = \frac{\mathbf{M}\{\xi_i^3\}}{6 \mathbf{D}^3\{\xi_i\}}.$$

As clearly

$$(7) \quad \mathbf{P}\{a_1 \leq \eta_n < b_1 | a_2 \leq \eta_n < b_2\} = \frac{\mathbf{P}\{a_1 \leq \eta_n < b_1\}}{\mathbf{P}\{a_2 \leq \eta_n < b_2\}} = \frac{F_n\left(\frac{b_1}{\sqrt{n}}\right) - F_n\left(\frac{a_1}{\sqrt{n}}\right)}{F_n\left(\frac{b_2}{\sqrt{n}}\right) - F_n\left(\frac{a_2}{\sqrt{n}}\right)}$$

and from (3) we obtain

$$\begin{aligned} F_n\left(\frac{b}{\sqrt{n}}\right) - F_n\left(\frac{a}{\sqrt{n}}\right) &= \Phi\left(\frac{b}{\sqrt{n}}\right) - \Phi\left(\frac{a}{\sqrt{n}}\right) + \frac{C}{\sqrt{n}} \left\{ \Phi'\left(\frac{b}{\sqrt{n}}\right) \left(1 - \frac{b^2}{n}\right) - \right. \\ &\quad \left. - \Phi'\left(\frac{a}{\sqrt{n}}\right) \left(1 - \frac{a^2}{n}\right) + o(1) \right\} = \frac{b-a}{\sqrt{n}} \left\{ \frac{\Phi\left(\frac{b}{\sqrt{n}}\right) - \Phi\left(\frac{a}{\sqrt{n}}\right)}{\frac{b}{\sqrt{n}} - \frac{a}{\sqrt{n}}} + o(1) \right\} = \\ &= \frac{b-a}{\sqrt{n}} \{ \Phi'(0) + o(1) \} \end{aligned}$$

it follows that

$$\mathbf{P}\{a_1 \leq \eta_n < b_1 | a_2 \leq \eta_n < b_2\} = \frac{b_1 - a_1 + o(1)}{b_2 - a_2 + o(1)},$$

which is equivalent to (1).

A. RÉNYI proved in [5] that the assumption (2) in the above sufficient condition may be replaced by the supposition that the variables ξ_i have a quadratically integrable density function.

We shall prove that the condition (2) may be replaced by the following condition: for the common characteristic function $\varphi(t)$ of the variables ξ_i we have

$$(8) \quad \lim_{|t| \rightarrow \infty} \sup |\varphi(t)| < 1.$$

Thus we prove the following

Theorem. Let ξ_n ($n = 1, 2, \dots$) be independent and identically distributed random variables having 0 mean and variance 1. Let $\varphi(t)$ denote the characteristic function of the distribution of the variables ξ_i and suppose that (8) holds. Then (1) is valid.

Remark. A sufficient condition ensuring the validity of (8) is that in the canonical decomposition of the distribution function $G(x)$ of the ξ_i -s

the absolute continuous component should have a positive weight (see [1], Theorem 7. p. 25.).

Proof of the theorem. If χ_j denotes the indicator function of the interval $I_j = (a_j, b_j)$, i. e.

$$(9) \quad \chi_j(x) = \begin{cases} 1, & \text{if } x \in I_j \\ 0 & \text{otherwise} \end{cases} \quad j = 1, 2$$

and $G_n(x)$ denotes the distribution function of

$$\eta_n = \sum_{i=1}^n \xi_i$$

we may express the conditional probability in (1) in the following form :

$$(10) \quad \mathbf{P} \{a_1 \leq \eta_n < b_1 \mid a_2 \leq \eta_n < b_2\} = \frac{\int_{-\infty}^{\infty} \chi_1(x) dG_n(x)}{\int_{-\infty}^{\infty} \chi_2(x) dG_n(x)}.$$

Since the Fourier—Stieltjes transforms of the functions $\chi_j(x)$ are not absolutely integrable, we cannot immediately insert the characteristic functions of the variables ξ_i . Therefore we have to apply certain approximations.

Let $\varepsilon > 0$ be an arbitrarily small number, for which

$$(11) \quad 0 < \varepsilon < \min \left(\frac{1}{4}, b_1 - a_1 \right).$$

Let \underline{h}_j resp. \bar{h}_j be lower resp. upper approximatory functions to the indicator functions χ_j which are twice continuously derivable, further \underline{h}_j is identical with χ_j outside the intervals $(a_j, a_j + \varepsilon)$ and $(b_j - \varepsilon, b_j)$ and analogously \bar{h}_j identical with χ_j outside the intervals $(a_j - \varepsilon, a_j)$ and $(b_j, b_j + \varepsilon)$, on the other hand these functions are monotonous in their two exceptional intervals. It is obvious from the construction that

$$(12) \quad \underline{h}_j(x) \leq \chi_j(x) \leq \bar{h}_j(x)$$

but

$$(13) \quad \int_{-\infty}^{\infty} \chi_j(x) dx - 2\varepsilon < \int_{-\infty}^{\infty} \bar{h}_j(x) dx - 2\varepsilon < \int_{-\infty}^{\infty} \chi_j(x) dx = b_j - a_j < \\ < \int_{-\infty}^{\infty} \underline{h}_j(x) dx + 2\varepsilon < \int_{-\infty}^{\infty} \chi_j(x) dx + 2\varepsilon \quad (j = 1, 2).$$

Thus according to (10)

$$(14) \quad \frac{\int_{-\infty}^{\infty} \underline{h}_1(x) dG_n(x)}{\int_{-\infty}^{\infty} \bar{h}_2(x) dG_n(x)} \leq \mathbf{P}\{a_1 \leq \eta_n < b_1 | a_2 \leq \eta_n < b_2\} \leq \frac{\int_{-\infty}^{\infty} \bar{h}_1(x) dG_n(x)}{\int_{-\infty}^{\infty} \underline{h}_2(x) dG_n(x)}.$$

Let $\underline{\psi}_j(t)$ resp. $\bar{\psi}_j(t)$ denote the Fourier transform of $\underline{h}_j(x)$ resp. $\bar{h}_j(x)$ ($j = 1, 2$) and $\varphi(t)$ the common characteristic function of the random variables ξ_i , i. e. put

$$(15) \quad \underline{\psi}_j(t) = \int_{-\infty}^{\infty} \underline{h}_j(x) e^{ixt} dx, \quad \bar{\psi}_j(t) = \int_{-\infty}^{\infty} \bar{h}_j(x) e^{ixt} dx, \quad \varphi(t) = \int_{-\infty}^{\infty} e^{itx} dG(x) \quad j=1, 2.$$

If $h(x)$ denotes any of $\underline{h}_j(x)$ and $\bar{h}_j(x)$ ($j = 1, 2$), while $\psi(t)$ denotes its Fourier transform, we obtain

$$(16) \quad \begin{aligned} \int_{-\infty}^{\infty} h(x) dG_n(x) &= \int_{-\infty}^{\infty} \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ixt} \psi(t) \right\} dG_n(x) = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(t) \left[\int_{-\infty}^{\infty} e^{-ixt} dG_n(x) \right] dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(t) \varphi(-t)^n dt. \end{aligned}$$

Applying this formula to the above functions, the integrals in (14) may be expressed by $\varphi(t)$; thus

$$(17) \quad \frac{\int_{-\infty}^{\infty} \underline{\psi}_1(t) \varphi(-t)^n dt}{\int_{-\infty}^{\infty} \underline{\psi}_2(t) \varphi(-t)^n dt} \leq \mathbf{P}\{a_1 \leq \eta_n < b_1 | a_2 \leq \eta_n < b_2\} \leq \frac{\int_{-\infty}^{\infty} \bar{\psi}_1(t) \varphi(-t)^n dt}{\int_{-\infty}^{\infty} \bar{\psi}_2(t) \varphi(-t)^n dt}.$$

Each operation applied in (16) is permissible. In order to justify the applicability of Fubini's theorem we have only to consider that the Fourier transforms occurring in (16) are absolutely integrable. But this follows from the fact that these functions are of order t^{-2} as it is easily seen by repeated partial integration.

In what follows we are going to demonstrate by the aid of Laplace's method that the quotients occurring in the inequality (17) converge for $n \rightarrow \infty$ to such limits which may be brought arbitrarily near to the quotient $(b_1 - a_1)/(b_2 - a_2)$ occurring in (1), by choosing ε sufficiently small.

We are going to prove by LAPLACE's method that

$$(18) \quad \sqrt{n} \int_{-\infty}^{\infty} \psi(t) \varphi(-t)^n dt = \sqrt{2\pi} \{ \psi(0) + \varepsilon O(1) + o(1) \}$$

where $\psi(t)$ may be any of the functions $\psi_1, \bar{\psi}_1, \psi_2, \bar{\psi}_2$.

To prove this let us choose the number $\delta > 0$ (depending on ε) so that for $|t| < \delta$

$$(19) \quad |\psi(t) - \psi(0)| < \varepsilon \quad \text{and} \quad |\varphi(t)| > \frac{1}{2}$$

should be fulfilled, further, that we should have

$$(20) \quad \log \varphi(-t) = -\frac{t^2}{2} + t^2 r(t)$$

with

$$(21) \quad |r(t)| < \varepsilon.$$

This is possible owing to the continuity of $\varphi(t)$ and to the relations $\varphi(0) = 1$, $\varphi'(0) = 0$ and $\varphi''(0) = -1$.

Considering that — according to (8) — the number

$$(22) \quad q_0 = \lim_{|t| \rightarrow \infty} \sup |\varphi(t)|$$

is less than 1, a number $T > \delta$ can be given for any q_1 between q_0 and 1 ($q_0 < q_1 < 1$) for which

$$(23) \quad |\varphi(t)| < q_1 \quad \text{for} \quad |t| > T.$$

It follows from (8) that $G(x)$ is a nonlattice distribution function; and therefore,

$$(24) \quad |\varphi(t)| < 1 \quad \text{for} \quad t \neq 0$$

(see [2], p. 59.); thus the maximum q_2 of the continuous function $|\varphi(t)|$ on the closed set $[-T, -\delta] \cup [\delta, T]$ is also less than 1. Putting $q = \max(q_1, q_2)$ obviously

$$(25) \quad |\varphi(t)| < q < 1 \quad \text{for} \quad |t| > \delta$$

and thus

$$(26) \quad \int_{-\infty}^{\infty} \psi(t) \varphi^n(-t) dt = \int_{-\delta}^{\delta} \psi(t) \varphi^n(-t) dt + R,$$

² $O(1)$ resp. $o(1)$ denote values depending on n which — ε being fixed — remain below an absolute constant resp. they converge to zero as $n \rightarrow \infty$.

where

$$(27) \quad |R| < q^n \left\{ \int_{-\infty}^{-\delta} |\psi(t)| dt + \int_{+\delta}^{\infty} |\psi(t)| dt \right\} = O(q^n) .$$

As, owing to $0 < q < 1$

$$(28) \quad \lim_{n \rightarrow \infty} q^n \sqrt{n} = 0 ,$$

it follows that

$$(29) \quad \sqrt{n} \int_{-\infty}^{\infty} \psi(t) \varphi^n(-t) dt = \sqrt{n} \left\{ \int_{-\delta}^{\delta} \psi(t) \varphi^n(-t) dt + o(1) \right\} .$$

As the second step of the proof we are going to apply (19) and (20) to estimate the integral on the right hand side of (29). First, according to (19), we obtain

$$(30) \quad \begin{aligned} \int_{-\delta}^{\delta} \psi(t) \varphi^n(-t) dt &= \psi(0) \int_{-\delta}^{\delta} e^{n \log \varphi(-t)} dt + \int_{-\delta}^{\delta} [\psi(t) - \psi(0)] e^{n \log \varphi(-t)} dt = \\ &= [\psi(0) + \varepsilon O(1)] \int_{-\delta}^{\delta} e^{n \log \varphi(-t)} dt . \end{aligned}$$

Applying (20) we may use the fact that

$$(31) \quad e^z = 1 + z g(z)$$

where

$$(32) \quad |g(z)| = \left| \frac{e^z - 1}{z} \right| < e^{|z|} .$$

It follows therefore that

$$(33) \quad \begin{aligned} \int_{-\delta}^{\delta} e^{n \log \varphi(-t)} dt &= \int_{-\delta}^{\delta} e^{-\frac{nt^2}{2}} \{1 + nt^2 r(t) g(nt^2 r(t))\} dt = \\ &= \int_{-\delta}^{\delta} e^{-\frac{nt^2}{2}} dt + n \int_{-\delta}^{\delta} r(t) t^2 e^{-\frac{nt^2}{2}} g(nt^2 r(t)) dt . \end{aligned}$$

For the first integral on the right side of (33)

$$(34) \quad \int_{-\delta}^{\delta} e^{-\frac{nt^2}{2}} dt = \frac{1}{\sqrt{n}} \int_{-\delta\sqrt{n}}^{\delta\sqrt{n}} e^{-\frac{u^2}{2}} du = \frac{\sqrt{2\pi}}{\sqrt{n}} \left\{ 1 - \frac{2}{\sqrt{2\pi}} \int_{\delta\sqrt{n}}^{\infty} e^{-\frac{u^2}{2}} du \right\} = \frac{\sqrt{2\pi}}{\sqrt{n}} [1 + o(1)]$$

and for the second

$$\begin{aligned} \left| n \int_{-\delta}^{\delta} r(t) t^2 e^{-\frac{nt^2}{2}} g(nt^2 r(t)) dt \right| &< \varepsilon n \int_{-\delta}^{\delta} t^2 e^{-\frac{nt^2}{2} + \varepsilon nt^2} dt = \\ &= \frac{\varepsilon}{\sqrt{n(1-2\varepsilon)^{3/2}}} \int_{-\delta\sqrt{n(1-2\varepsilon)}}^{\delta\sqrt{n(1-2\varepsilon)}} u^2 e^{-\frac{u^2}{2}} du < \frac{\sqrt{2\pi}\varepsilon}{\sqrt{n}} 2\sqrt{2}. \end{aligned}$$

Finally

$$(35) \quad \int_{-\delta}^{\delta} e^{n \log \varphi(-t)} dt = \frac{\sqrt{2\pi}}{\sqrt{n}} \{1 + \varepsilon O(1) + o(1)\}.$$

Inserting this into (30) we obtain

$$\int_{-\delta}^{\delta} \psi(t) \varphi^n(-t) dt = \frac{\sqrt{2\pi}}{\sqrt{n}} \{\psi(0) + \varepsilon O(1) + o(1)\}$$

and — according to (29) — this is the same as (18).

Let us use (18) to estimate (16). Since, according to (13), for lines both under and above

$$(36) \quad \psi_j(0) = \int_{-\infty}^{\infty} h_j(x) dx = b_j - a_j + \varepsilon O(1),$$

we obtain the same limiting value $(b_1 - a_1)/(b_2 - a_2)$, after substituting and first taking $n \rightarrow \infty$, then $\varepsilon \rightarrow 0$.

Thus our Theorem is proved.

Acknowledgement. I wish to express my thanks to A. RÉNYI and to T. LIPTÁK for their valuable critical remarks.

(Received 8 July, 1958.)

REFERENCES

- [1] CRAMÉR, H.: *Random variables and probability distributions*. Cambridge Tracts in Mathematics, No. 36., Cambridge, 1937.
- [2] GNEDENKO, B. V.—KOLMOGOROV, A. N.: *Limit distributions for sums of independent random variables*. Addison—Wesley, Cambridge (Mass.), 1954.
- [3] KALLIANPUR, G.—ROBBINS, H.: „The sequence of sums of independent random variables”. *Duke Mathematical Journal* **21** (1954) 285—307.
- [4] RÉNYI, A.: „On a new axiomatic theory of probability”. *Acta Mathematica Academiae Scientiarum Hungaricae* **6** (1955) 315.

EGY FELTÉTELES HATÁRELOSZLÁSTÉTELI ŐL

SCHNELL LÁSZLÓNÉ

Kivonat

A szerző a következő tételt bizonyítja be : *ha a független, azonos eloszlású, 0 várható értékű és véges szórású ξ_i valószínűségi változók közös $\varphi(t)$ karakterisztikus függvényére*

$$(1) \quad \limsup_{|t| \rightarrow \infty} |\varphi(t)| < 1,$$

akkor fennáll a következő feltételes határeloszlástétel

$$(2) \quad \lim_{n \rightarrow \infty} \mathbf{P} \left\{ a_1 \leq \sum_{i=1}^n \xi_i < b_1 \mid a_2 \leq \sum_{i=1}^n \xi_i < b_2 \right\} = \frac{b_1 - a_1}{b_2 - a_2}$$

ha $a_2 \leq a_1 < b_1 \leq b_2$. A bizonyítás a Laplace-féle módszeren alapul. Az (1) feltétel teljesül például akkor, ha a ξ_i változók közös $G(x)$ eloszlásfüggvényének kanonikus felbontásában pozitív súllyal szerepel az abszolút folytonos komponens.

ОБ ОДНОЙ ТЕОРЕМЕ ОБ УСЛОВНОМ ПРЕДЕЛЬНОМ РАСПРЕДЕЛЕНИИ ДЛЯ ОДИНАКОГО РАСПРЕДЕЛЕННЫХ НЕЗАВИСИМЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

E. SCHNELL

Резюме

Автор доказывает следующую теорему : если для общей характеристической функций $\varphi(t)$ независимых одинаково распределенных случайных величин ξ_i с нулевым математическим ожиданием и конечной дисперсией имеет место соотношение

$$(1) \quad \limsup_{|t| \rightarrow \infty} |\varphi(t)| < 1$$

то имеет место следующая теорема о предельном распределении :

$$(2) \quad \lim_{n \rightarrow \infty} \mathbf{P} \left\{ a_1 \leq \sum_{i=1}^n \xi_i < b_1 \mid a_2 \leq \sum_{i=1}^n \xi_i < b_2 \right\} = \frac{b_1 - a_1}{b_2 - a_2}$$

если $a_2 \leq a_1 < b_1 \leq b_2$. Доказательство основывается на методе Laplace-a. Условие (1) выполняется, например, в том случае, если в каноническом разложении общей функции распределения с положительным весом фигурирует абсолютно непрерывная компонента.

VÁLLALATI ÉS NÉPGAZDASÁGI MEGTAKARÍTÁS

BRÓDY ANDRÁS¹

1. §. Bevezetés

Tegyük fel, hogy az árrendszer *értékarányos*, amin azt értjük, hogy az egyes termékek ára arányos a termékekben megtestesülő összes közvetlen és közvetett munkaráfordítással. Az arányossági tényezőt 1-nek tételezve röviden úgy tekintjük, hogy az árak munkaórákban vannak megadva. Ezt a következőképpen értjük: Legyen adva N számú vállalat, mindegyikük előállít egy bizonyos „ k ” fajta terméket ($k = 1, 2, \dots, N$). A k sorszámú vállalatnál ($k = 1, 2, \dots, N$) a termék egységének² (tehát egy egységnyi „ k ” terméknek) előállításához az „ i ” termékéből a_{ik} egységre van szükség ($i = 1, 2, \dots, N$), továbbá még m_k óra munkára. Az egyes termékek egységében megtestesülő összes munkamennyiséget, tehát a közvetlenül ráfordított m_k munkaóra és a többi termék felhasználása révén közvetve ráfordított munkaórák összegét (a termékek „árát”) a W. LEONTIEF [1] által felállított lineáris egyenletrendszer segítségével fejezhetjük ki.³

Ezen egyenletrendszer matrix írásmóddal való felírásához vezessük be az alábbi jelöléseket:

$$(1) \quad \mathbf{m}^* = [m_1, m_2, \dots, m_N]$$

$$(2) \quad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{bmatrix}$$

$$(3) \quad \mathbf{p}^* = [p_1, p_2, \dots, p_N],$$

ahol $p_i > 0$ jelenti az „ i ” termék egységének munkaórában kifejezett árát ($i = 1, 2, \dots, N$). \mathbf{A} az úgynevezett technológiai matrix. A LEONTIEF-féle egyenletrendszert akkor a következő alakba lehet írni:

$$(4) \quad \mathbf{p}^* \mathbf{A} + \mathbf{m}^* = \mathbf{p}^*,$$

¹ Közgazdaságtudományi Intézet.

² Az egységek megválasztása tárgyalásunk szempontjából mellékes; lehet darab, liter stb.

³ Az ilyen típusú egyenletrendszer valamivel részletesebb tárgyalását magyar nyelven lásd: [2], [3].

amiből az

$$(5) \quad (\mathbf{E} - \mathbf{A})^{-1} = \mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ r_{N1} & r_{N2} & \dots & r_{NN} \end{bmatrix}$$

jelöléssel (\mathbf{E} az N -edrendű egységmatrix)

$$(6) \quad \mathbf{m}^* \mathbf{R} = \mathbf{p}^* .$$

Hasonló összefüggések vannak az egyes vállalatok termelésének arányai között is. Ha az $\mathbf{x}^* = [x_1, x_2, \dots, x_N]$ vektor mutatja, hogy az egyes termékekből hány egységet állítanak elő adott időszakban, akkor az \mathbf{Ax} vektor mutatja, hogy ezek előállításához ugyanezen termékekből milyen mennyiségeket kell felhasználni, a végső fogyasztásra rendelkezésre álló mennyiségeket pedig az $\mathbf{y} = \mathbf{x} - \mathbf{Ax}$ vektor adja meg. Az $\mathbf{y}^* = [y_1, y_2, \dots, y_N]$ vektor jellemzi tehát az egyes termékekből a termelési időszak végén a társadalom rendelkezésére álló végtermékek mennyiségét. Ezek a végtermékek képezik a termelési folyamat tulajdonképpeni célját. Az \mathbf{x} és \mathbf{y} vektorok között tehát az alábbi összefüggés áll fenn:

$$(7) \quad \mathbf{x} = \mathbf{R} \mathbf{y} .$$

Tegyük fel mármost, hogy például a „ k ” termék egységének előállításához felhasznált a_{ik} egységnyi „ i ” termék helyett a következőkben már csak $a_{ik} - h$ ($0 < h \leq a_{ik}$) mennyiségű „ i ” terméket fogunk felhasználni, azaz minden egység előállításánál ezentúl megtakarítunk h egységnyi „ i ” terméket. Kérdés, hogy ez a megtakarítás mekkorának mutatkozik az árrendszer segítségével számolva az egyes vállalatok számára, és mekkora a valóságban, a tényleges társadalmi munkaráfordítást figyelembe véve?

2. §. Vállalati és népgazdasági megtakarítás

A vállalat számára a számítás egyszerű: h egységnyi „ i ” terméket takarít meg a „ k ” termék termelésének minden egysége, összesen tehát x_k egység után. Az „ i ” termék egységára p_i , az összes számított megtakarítás tehát

$$(8) \quad \Delta_1 = h p_i x_k .$$

Népgazdasági szinten a számítás jóval bonyolultabb: ez annak a következménye, hogy a megtakarítás maga után vonja a technológiai matrix és ezzel az egész árrendszer és a termelés arányainak megváltozását.

A technológiai változás előtt a társadalom összes ráfordítása valamely \mathbf{y} végső fogyasztás biztosítása érdekében $\mathbf{m}^* \mathbf{R} \mathbf{y}$ volt. (E kifejezést kétféleképp is értelmezhetjük: $\mathbf{m}^* \mathbf{x}$ az egyes termékekből termelt egységek számának szorzata az egységekre eső munkaráfordításokkal, $\mathbf{p}^* \mathbf{y}$ formájában pedig a végső fogyasztásra kerülő termékek és munkaórákban kifejezett áruk szorzata.) A megváltozott technológiai matrix miatt a társadalom új munkaráfordítása azonos végső fogyasztás biztosítására $\mathbf{m} \mathbf{R}_h \mathbf{y}$ lesz, ahol

$$(9) \quad \mathbf{R}_h = \mathbf{R} - \frac{h}{1 + h r_{ki}} \mathbf{R} \mathbf{E}_{ik} \mathbf{R}$$

(lásd: [4], (3.25), 33. oldal). E_{ik} itt olyan N -edrendű kvadratikus matrixot jelent, amelyben az i -edik sor k -adik oszlopában 1 áll, a többi helyen pedig 0. A tényleges társadalmi megtakarítás tehát:

$$(10) \quad \Delta_2 = m^* R y - m^* R_h y = \frac{h}{1 + hr_{ki}} m^* R E_{ik} R y.$$

A (6) és (7) relációk figyelembevételével tehát

$$(11) \quad \Delta_2 = \frac{h}{1 + hr_{ki}} p^* E_{ik} x = \frac{h}{1 + hr_{ki}} p_i x_k.$$

A tényleges népgazdasági megtakarítás aránya tehát a vállalatnál számított megtakarításhoz képest (8) és (11) alapján a következő:

$$(12) \quad \frac{\Delta_1}{\Delta_2} = 1 + hr_{ki}.$$

Kimondhatjuk tehát: 1. A tényleges népgazdasági megtakarítás mindig kisebb, mint a vállalat által számított. 2. Népgazdasági szempontból nem mindegy melyik iparágban vagy vállalatnál érnek el valamely megtakarítást vagy a vállalaton belül melyik termék felhasználásában érnek el megtakarítást, még akkor sem, ha a közvetlen vállalati szinten számított megtakarítás azonosnak látszik is.

Meg kell jegyezni, hogy az eltérés a kétféle megtakarítás közt általában nem jelentős. Az eltérés gyakorlati adatok alapján számítva a legtöbb esetben néhány ezreléket, legfeljebb 1—2 százalékot tesz ki. Mégis speciális helyzetekben (így elsősorban az egyes termékek önfogyasztása, a_{ii} esetében) a kétfajta megtakarítás nagysága jelentősebben is eltérhet egymástól.

(Beérkezett: 1958. szeptember 15.)

IRODALOM

- [1] LEONTIEF, W. W.: *The Structure of American Economy 1919—1939*. Oxford University Press, Oxford 1953. (2nd ed.)
- [2] RÉNYI, A — BRÓDY, A.: „Az árrendezés problémájáról”. *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* 1 (1956) 325—335.
- [3] BRÓDY, A.: „Az ágazati kapcsolatok mérlegének elméletéhez.” *Közgazdasági Szemle*. 8—9 (1958) 823—835.
- [4] BODEWIG, E.: *Matrix calculus*. North-Holland Publishing Company, Amsterdam, 1956.

ОБ ЭКОНОМИИ ПРЕДПРИЯТИЯ И НАРОДНОГО ХОЗЯЙСТВА

A. BRÓDY

Резюме

Вычисленная с помощью цен экономия труда (Δ_1) не тождественна действительной, народнохозяйственной экономии труда (Δ_2), даже в том случае, когда цены пропорциональны ценностям.

Если любой элемент a_{ik} технологической матрицы Леонтьева $A = [a_{ik}]$ уменьшить на величину h , то

$$\Delta_1 = hp_i x_k,$$

где p_i цена продукта « i », выраженная числом рабочих часов, а x_k количество созданного продукта « k », в то время как

$$\Delta_2 = \frac{\Delta_1}{1 + hr_{ki}}$$

где r_{ki} соответствующий элемент матрицы $R = (E - A)^{-1}$.

SAVINGS, COMPUTED ON A PRICE BASIS AND SAVINGS COMPUTED ON THE BASIS OF A TECHNOLOGICAL MATRIX

by

A. BRÓDY

Abstract

The volume of labour economised when computed on a price basis (Δ_1) is not equal to the actual, social savings (Δ_2), even if all prices were proportional to the respective commodity-values.

Assuming a decrease of h in any element a_{ik} of a LEONTIEF-type technological matrix,

$$\Delta_1 = hp_i x_k$$

where p_i stands for the price of the i -th commodity, measured by work hours, and x_k stands for the quantity produced of the k -th commodity; whereas

$$\Delta_2 = \frac{1}{1 + hr_{ki}} \Delta_1$$

where r_{ki} is the corresponding element of the LEONTIEF-inverse

$$R = (E - A)^{-1}.$$

KOMBINATORIKUS MÓDSZER A SZÁLLÍTÁSI PROBLÉMA MEGOLDÁSÁRA

EGERVÁRY JENŐ

A lineáris programozás egyik alapfeladata, a szállításprobléma, a következő minimumfeladat megoldását teszi szükségessé. Adva van k számú termelő hely, melyek ugyanazt a terméket állítják elő, és pedig a

$$\begin{aligned} T_1 \text{ termelő hely tárol } v_1 \text{ egységet} \\ T_2 \text{ termelő hely tárol } v_2 \text{ egységet} \\ \dots\dots\dots \\ T_k \text{ termelő hely tárol } v_k \text{ egységet.} \end{aligned}$$

Adva van továbbá l fogyasztó hely, melyek a fenti termékeket igénylik és pedig az

$$\begin{aligned} F_1 \text{ fogyasztó hely igényel } \mu_1 \text{ egységet} \\ F_2 \text{ fogyasztó hely igényel } \mu_2 \text{ egységet} \\ \dots\dots\dots \\ F_l \text{ fogyasztó hely igényel } \mu_l \text{ egységet.} \end{aligned}$$

Először az egyszerűsítés érdekében feltesszük, hogy a termelő helyeken tárolt összes termék egyenlő a fogyasztó helyek összes szükségletével, azaz

$$(1) \quad v_1 + v_2 + \dots + v_k = \mu_1 + \mu_2 + \dots + \mu_l$$

Legyen továbbá egy egység szállítási költsége a T_i termelő helyről az F_j fogyasztó helyre c_{ij} ($i = 1, 2, \dots, k$; $j = 1, 2, \dots, l$), és tegyük fel, hogy ezek a szállítási költségek nem-negatív egész számok.

Kérdés, hogy kell a szállítást optimálisan megszervezni, vagyis a szállítási összköltséget minimumra redukálni.

Legyen a T_i termelési helyről az F_j fogyasztási helyre szállítandó egységek száma x_{ij} ($i = 1, 2, \dots, k$; $j = 1, 2, \dots, l$). Feltetés szerint a T_i termelési hely v_i egységet tárol, tehát a T_i termelési helyről elszállított termékek összegére fenn kell állnia az

$$(2) \quad x_{i1} + x_{i2} + \dots + x_{il} = v_i$$

egyenletnek ($i = 1, 2, \dots, k$).

Ugyancsak feltetés szerint az F_j fogyasztási hely μ_j egységet igényel, tehát az F_j fogyasztási helyre szállítandó termékek összegére fenn kell állnia az

$$(3) \quad x_{1j} + x_{2j} + \dots + x_{kj} = \mu_j$$

egyenletnek ($j = 1, 2, \dots, l$).

A szállítás összköltsége a fentiek szerint nyilván

$$(4) \quad \sum_{i=1}^k \sum_{j=1}^l c_{ij} x_{ij}.$$

A szállítási probléma matematikai megfogalmazásban tehát azt jelenti, hogy a (4) lineáris forma minimumát kell meghatározni a (2) és (3) feltételi egyenletek és — a probléma természetéből folyó — $x_{ij} \geq 0$ feltételi egyenlőtlenségek kielégítése mellett ($i = 1, 2, \dots, k$; $j = 1, 2, \dots, l$).

Hogy az alább kifejtendő módszer lényegét szemléletesen bemutathassuk, foglalkozzunk először a szállítási problémának azzal az egyszerű speciális esetével, melyben a termelő és fogyasztási helyek száma megegyezik, továbbá mindegyik termelő hely egy egységet tárol és mindegyik fogyasztó hely egy egységet igényel. Ebben az esetben a szállítási probléma ekvivalens az ún. hozzárendelési problémával („assignment problem”). Mindegyik termelő helyről ugyanis az ott tárolt egységet valamelyik fogyasztó helyre kell szállítani és a feladat éppen abban áll, hogy a

T_1 termelő helyhez egy F_{i_1} fogyasztó helyet,

T_2 termelő helyhez egy F_{i_2} fogyasztó helyet,

.....

T_n termelő helyhez egy F_{i_n} fogyasztó helyet

úgy rendeljünk hozzá, hogy a szállítás összköltségét szolgáltató

$$(5) \quad c_{1i_1} + c_{2i_2} + \dots + c_{ni_n}$$

„permutációs összeg” minimális legyen. Nyilván a különböző permutációs összegek száma $n!$ és a szállítási probléma megoldása céljából ezen összegek közül a legkisebbet kell kikeresnünk.

A keresett minimum egy (valóságos szállítási problémánál elő nem forduló) esetben triviális módon evidenssé válik, ti. akkor, ha a nem-negatív c_{ij} elemekből álló n -edrendű kvadratikus matrix tartalmaz n „független 0-elemet”, vagyis n olyan 0-t, melyek közül nincs kettő ugyanabban a sorban vagy oszlopban. Ekkor ugyanis ezek a 0-ok egy olyan permutációs összeget alkotnak, mely nyilván kisebb (vagy legalább nem nagyobb), mint az összes többi permutációs összeg. Meg fogjuk mutatni, hogy egy alkalmas transzformációval mely a permutációs összegek nagysági sorrendjét változtatlanul hagyja, nem-negatív elemű kvadratikus matrix olyan matrixba transzformálható, mely n független 0-elemet tartalmaz, melyre nézve tehát a minimum-feladat megoldása triviális.

Tekintsük először a $\gamma_{ij} = u_i + v_j$ elemekből álló

$$(6) \quad \begin{bmatrix} u_1 + v_1 & u_1 + v_2 & \dots & u_1 + v_n \\ u_2 + v_1 & u_2 + v_2 & \dots & u_2 + v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_n + v_1 & u_n + v_2 & \dots & u_n + v_n \end{bmatrix}.$$

n -edrendű kvadratikus matrixot. Az ezen matrixból képezhető

$$(7) \quad \begin{aligned} \gamma_{1i_1} + \gamma_{2i_2} + \dots + \gamma_{ni_n} = \\ = u_1 + v_{i_1} + u_2 + v_{i_2} + \dots + u_n + v_{i_n} = \sum_{i=1}^n u_i + \sum_{j=1}^n v_j \end{aligned}$$

permutációs összegek nyilván mind egyenlők egymással. Ha tehát a kiindulási $[c_{ij}]$ matrixból a

$$(8) \quad [c'_{ij}] = [c_{ij} - \gamma_{ij}] = [c_{ij} - u_i - v_j]$$

matrixot képezzük, akkor a transzformált $[c'_{ij}]$ matrixhoz tartozó

$$(9) \quad c'_{1i_1} + c'_{2i_2} + \dots + c'_{ni_n}$$

permutációs összegek mindegyike a

$$(10) \quad \sum_{i=1}^n u_i + \sum_{j=1}^n v_j$$

mennyiséggel lesz kisebb, mint az eredeti

$$(11) \quad c_{1i_1} + c_{2i_2} + \dots + c_{ni_n}$$

permutációs összegek. A (8) transzformáció tehát a permutációs összegek nagysági sorrendjét változtatlanul hagyja. Ha tehát tudunk olyan (8) alakú transzformációt találni, melynek végrehajtása után a transzformált $[c'_{ij}]$ matrix n független 0-elemt tartalmaz és továbbra is nem-negatív elemekkel bír, akkor a szállítási probléma most tárgyalt speciális esetét megoldottuk, mert a független 0-ok megadják a termelési helyeknek és a fogyasztási helyeknek azt (vagy legalábbis egy olyan) egymáshoz rendelését, melynél a szállítások összköltsége minimális.

Például az

$$(12) \quad \begin{bmatrix} 1 & 9 & 5 & 8 \\ 2 & 3 & 0 & 4 \\ 3 & 1 & 4 & 1 \\ 2 & 7 & 2 & 8 \end{bmatrix}$$

negyedrendű matrix esetén az

$$(13) \quad u_1 = 1; u_2 = 0; u_3 = -2; u_4 = 2; v_1 = 0; v_2 = 3; v_3 = 0; v_4 = 3$$

értékekkel végrehajtott (8) típusú transzformáció a következő eredményt adja:

$$(14) \quad \begin{bmatrix} 1 & 9 & 5 & 8 \\ 2 & 3 & 0 & 4 \\ 3 & 1 & 4 & 1 \\ 2 & 7 & 2 & 8 \end{bmatrix} - \begin{bmatrix} 1+0 & 1+3 & 1+0 & 1+3 \\ 0+0 & 0+3 & 0+0 & 0+3 \\ -2+0 & -2+3 & -2+0 & -2+3 \\ 2+0 & 2+3 & 2+0 & 2+3 \end{bmatrix} = \begin{bmatrix} 0 & 5 & 4 & 4 \\ 2 & 0 & 0 & 1 \\ 5 & 0 & 6 & 0 \\ 0 & 2 & 0 & 3 \end{bmatrix}.$$

Az így nyert negyedrendű transzformált matrix már tartalmaz négy független 0-elemt (lásd a vastagon szedett 0-okat!) és az ezeknek megfelelő

$$1 + 3 + 1 + 2 = 7$$

permutációs összeg a keresett minimum.

Az alábbiakban kimutatjuk, hogy a kívánt tulajdonságú transzformált matrixhoz véges számú iteratív lépésben mindig eljuthatunk. Azt, hogy a transzformált matrix minden sora és minden oszlopa legalább egy 0-elemet tartalmazzon és amellet az összes elemek nem-negatívok maradjanak, elérhetjük triviális módon:

1° Vonjuk ki mindegyik sor elemeiből az illető sor legkisebb elemét. A fenti példa esetén adódik:

$$(15) \quad \begin{bmatrix} 1 & 9 & 5 & 8 \\ 2 & 3 & 0 & 4 \\ 3 & 1 & 4 & 1 \\ 2 & 7 & 2 & 8 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 8 & 4 & 7 \\ 2 & 3 & 0 & 4 \\ 2 & 0 & 3 & 0 \\ 0 & 5 & 0 & 6 \end{bmatrix}.$$

2° Vonjuk ki mindegyik oszlop elemeiből az illető oszlop legkisebb elemét. A fenti példa esetén ez a transzformáció a (15) jobboldalán szereplő matrixot változatlanul hagyja. (Nyilván az 1° és 2° transzformációk (8)-nak speciális esetei).

Az így transzformált matrix nyilván legalább n 0-elemet tartalmaz, de még nem tartalmaz szükségképpen n független 0-elemet.

A további (8) típusú transzformációk, melyek véges számú lépésben n független 0-elemet tartalmazó és nem-negatív elemekkel bíró transzformált matrixot eredményeznek, egy kombinatorikus tételre alapulnak, melyet jelen dolgozat szerzője 1931-ben publikált [1]:

Ha egy kvadrátikus matrix elemei nem-negatív egész számok, akkor az olyan vonalak (sorok vagy (és) oszlopok) minimális száma, melyek az összes 0-elemeket tartalmazzák, egyenlő a matrixban található független 0-elemek maximális számával.

Szemléletesebb fogalmazásban: Ha az n -edrendű matrixban található 0-elemek „lefedéséhez” legalább k ($\leq n$) vonal (sor, oszlop) szükséges, akkor ezen matrix elemei közt van k (de nem több) olyan 0-elem, melyek közül nincs kettő egy vonalban (sorban, oszlopban).

Tegyük fel mármost, hogy a fenti 1°, 2° triviális transzformációval nyert matrix — az általános esetnek megfelelően — nem tartalmaz n független 0-elemet és legyen a belőle kiválasztható független 0 elemek maximális száma $k < n$. Ekkor az idézett [1] tétel szerint az összes 0-elemek fedhetők k vonallal. Alkalmazzuk most a következő (8) típusú transzformációkat:

I° Vonjuk ki a matrix valamennyi eleméből a legkisebb *fedetlen* elemet. Ezáltal a matrix egyes elemei negatívvá válnak.

II° Ezeknek a negatív elemeknek a kiküszöbölése céljából adjuk hozzá a 0-elemeket fedő vonalak minden eleméhez a fenti legkisebb fedetlen elemet.

Az I°, II° lépésekből álló transzformációnál nyilván az összes elemek nem-negatívok maradnak.

Ha továbbá a legkisebb fedetlen elem ε volt, akkor az I° lépésnél a matrix-elemek összege $n^2 \varepsilon$ -nal csökken, a II° lépésnél pedig $n k \varepsilon$ -nal növekszik, tehát végeredményben a most leírt transzformációnál a matrixelemek összege $n(n - k) \varepsilon$ -nal csökken. Mindaddig tehát, amíg a transzformált matrix 0-elemei $k < n$ vonallal fedhetők, az elemek összege egy újabb transzformáció által

$n(n-k)\varepsilon$ -nal csökkenthető. Ily módon nyilván véges számú lépésben eljutunk egy olyan transzformált matrixhoz, melynek 0-elemei kevesebb, mint n vonallal nem fedhetők.

Ekkor azonban a fentidézett [1] tétel szerint a transzformált matrixban található n olyan 0 elem, melyek közül nincs kettő egy vonalban és ezek a 0-elemek a termelő és fogyasztó helyeknek olyan egymáshoz rendelését adják, melynél a szállítási összköltség minimális.

Például a fenti (12) matrix 0-elemei az első oszloppal, a harmadik oszloppal és a harmadik sorral, azaz összesen $3 (< 4)$ vonallal fedhetők (a nyíllal megjelölt oszlopok, illetve sorok mutatják a fedővonalakát):

$$(16) \quad \begin{array}{c} \downarrow \quad \downarrow \\ \rightarrow \begin{bmatrix} 0 & 8 & 4 & 7 \\ 2 & 3 & 0 & 4 \\ 2 & 0 & 3 & 0 \\ 0 & 5 & 0 & 6 \end{bmatrix} \end{array}.$$

A (16) matrix minden eleméből a legkisebb fedetlen elemnek, 3-nak a kivonásával adódik (17). Végül a fedővonalak minden eleméhez 3-at hozzáadva, nyerjük a transzformált (18) matrixot:

$$(17) \quad \begin{bmatrix} -3 & 5 & 1 & 4 \\ -1 & 0 & -3 & 1 \\ -1 & -3 & 0 & -3 \\ -3 & 2 & -3 & 3 \end{bmatrix}.$$

$$(18) \quad \begin{bmatrix} 0 & 5 & 4 & 4 \\ 2 & 0 & 0 & 1 \\ 5 & 0 & 6 & 0 \\ 0 & 2 & 0 & 3 \end{bmatrix}$$

Ennek a negyedrendű matrixnak az elemei már nem fedhetők kevesebb mint 4 vonallal, ez a matrix tehát az idézett tétel szerint szükségképpen tartalmaz 4 független 0-elemet. A vastagon szedett 0-ok függetlenek és az azoknak megfelelő permutációs összeg: $1 + 3 + 1 + 2$ valóban a már korábban megállapított minimumot szolgáltatja.

A szállítási probléma most tárgyalt speciális esetéről könnyen áttérhetünk az általános esetre. Tegyük fel, hogy

a T_1, \dots, T_{v_1} termelési helyeket egyesítjük egy v_1 egységet tároló \tilde{T}_1 ,

a $T_{v_1+1}, \dots, T_{v_1+v_2}$ termelési helyeket egyesítjük egy v_2 egységet tároló \tilde{T}_2 ,

.....

a T_{n-v_k+1}, \dots, T_n termelési helyeket egyesítjük egy v_k egységet tároló \tilde{T}_k termelési helyé. Továbbá

az F_1, \dots, F_{μ_1} fogyasztási helyeket egyesítjük egy μ_1 egységet igénylő \tilde{F}_1 ,

az $F_{\mu_1+1}, \dots, F_{\mu_1+\mu_2}$ fogyasztási helyeket egyesítjük egy μ_2 egységet igénylő \tilde{F}_2 ,

.....

az $F_{n-\mu_l+1}, \dots, F_n$ fogyasztási helyeket egyesítjük egy μ_l egységet igénylő \tilde{F}_l fogyasztási helyé.

Ekkor a

$$(19) \quad \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \dots & \mathbf{C}_{1l} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \dots & \mathbf{C}_{2l} \\ \dots & \dots & \dots & \dots \\ \mathbf{C}_{k1} & \mathbf{C}_{k2} & \dots & \mathbf{C}_{kl} \end{bmatrix}; \mathbf{C}_{ij} = \underbrace{\begin{bmatrix} c_{ij} & c_{ij} & \dots & c_{ij} \\ c_{ij} & c_{ij} & \dots & c_{ij} \\ \dots & \dots & \dots & \dots \\ c_{ij} & c_{ij} & \dots & c_{ij} \end{bmatrix}}_{\nu_i \text{ oszlop}} \left. \vphantom{\begin{bmatrix} c_{ij} & c_{ij} & \dots & c_{ij} \\ c_{ij} & c_{ij} & \dots & c_{ij} \\ \dots & \dots & \dots & \dots \\ c_{ij} & c_{ij} & \dots & c_{ij} \end{bmatrix}} \right\} \mu_j \text{ sor}$$

hipermatrixhoz tartozó hozzárendelési probléma nyilván ekvivalens a következő általános szállítási problémával:

$$(20) \quad \begin{array}{c|c|c|c|c|c} & \tilde{F}_1 & \tilde{F}_2 & & \tilde{F}_l & \\ \hline \tilde{T}_1 & c_{11} & c_{12} & \dots & c_{1l} & \leftarrow \nu_1 \\ \hline \tilde{T}_2 & c_{21} & c_{22} & \dots & c_{2l} & \leftarrow \nu_2 \\ \hline \vdots & \dots & \dots & \dots & \dots & \\ \hline \tilde{T}_k & c_{k1} & c_{k2} & \dots & c_{kl} & \leftarrow \nu_k \\ \hline & \downarrow & \downarrow & & \downarrow & \\ & \mu_1 & \mu_2 & & \mu_l & \end{array}$$

ahol $\nu_1, \nu_2, \dots, \nu_k$ illetve $\mu_1, \mu_2, \dots, \mu_l$ a $\tilde{T}_1, \tilde{T}_2, \dots, \tilde{T}_k$ termelési helyekről elszállítandó, illetve az $\tilde{F}_1, \tilde{F}_2, \dots, \tilde{F}_l$ fogyasztási helyekre szállítandó egységek számát jelentik és c_{ij} továbbra is az egység \tilde{T}_k termelési helyről \tilde{F}_l fogyasztási helyre való szállításának költsége.

Példaképpen bemutatjuk a következő szállítási problémát

$$(21) \quad \begin{array}{c|c|c|c|c|c|c} & \tilde{F}_1 & \tilde{F}_2 & \tilde{F}_3 & \tilde{F}_4 & \tilde{F}_5 & \\ \hline \tilde{T}_1 & 10 & 20 & 5 & 9 & 10 & \leftarrow 9 \\ \hline \tilde{T}_2 & 2 & 10 & 8 & 30 & 6 & \leftarrow 4 \\ \hline \tilde{T}_3 & 1 & 20 & 7 & 10 & 4 & \leftarrow 8 \\ \hline & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\ & 3 & 5 & 4 & 6 & 3 & \end{array}$$

amely az előbb mondottak szerint a következő hozzárendelési problémával ekvivalens:

(22)

10	10	10	20	20	20	20	20	5	5	5	5	9	9	9	9	9	9	10	10	10	10
10	10	10	20	20	20	20	20	5	5	5	5	9	9	9	9	9	9	10	10	10	10
10	10	10	20	20	20	20	20	5	5	5	5	9	9	9	9	9	9	10	10	10	10
10	10	10	20	20	20	20	20	5	5	5	5	9	9	9	9	9	9	10	10	10	10
10	10	10	20	20	20	20	20	5	5	5	5	9	9	9	9	9	9	10	10	10	10
10	10	10	20	20	20	20	20	5	5	5	5	9	9	9	9	9	9	10	10	10	10
10	10	10	20	20	20	20	20	5	5	5	5	9	9	9	9	9	9	10	10	10	10
10	10	10	20	20	20	20	20	5	5	5	5	9	9	9	9	9	9	10	10	10	10
10	10	10	20	20	20	20	20	5	5	5	5	9	9	9	9	9	9	10	10	10	10
2	2	2	10	10	10	10	10	8	8	8	8	30	30	30	30	30	30	6	6	6	6
2	2	2	10	10	10	10	10	8	8	8	8	30	30	30	30	30	30	6	6	6	6
2	2	2	10	10	10	10	10	8	8	8	8	30	30	30	30	30	30	6	6	6	6
2	2	2	10	10	10	10	10	8	8	8	8	30	30	30	30	30	30	6	6	6	6
1	1	1	20	20	20	20	20	7	7	7	7	10	10	10	10	10	10	4	4	4	4
1	1	1	20	20	20	20	20	7	7	7	7	10	10	10	10	10	10	4	4	4	4
1	1	1	20	20	20	20	20	7	7	7	7	10	10	10	10	10	10	4	4	4	4
1	1	1	20	20	20	20	20	7	7	7	7	10	10	10	10	10	10	4	4	4	4
1	1	1	20	20	20	20	20	7	7	7	7	10	10	10	10	10	10	4	4	4	4
1	1	1	20	20	20	20	20	7	7	7	7	10	10	10	10	10	10	4	4	4	4
1	1	1	20	20	20	20	20	7	7	7	7	10	10	10	10	10	10	4	4	4	4
1	1	1	20	20	20	20	20	7	7	7	7	10	10	10	10	10	10	4	4	4	4

$\mu_1 = 3$ $\mu_2 = 5$ $\mu_3 = 4$ $\mu_4 = 6$ $\mu_5 = 3$

$v_1 = 9$
 $v_2 = 4$
 $v_3 = 8$

A fentebb kifejtett transzformációs módszerrel nyerjük, hogy a szállítás összköltsége az ábrán kijelölt hozzárendelésnél minimális, azaz a

$$\text{Spur} \begin{bmatrix} 10 & 20 & 5 & 9 & 10 \\ 2 & 10 & 8 & 20 & 6 \\ 1 & 20 & 7 & 10 & 4 \end{bmatrix} \begin{bmatrix} x_{11} & x_{21} & x_{31} \\ x_{12} & x_{22} & x_{32} \\ x_{13} & x_{23} & x_{33} \\ x_{14} & x_{24} & x_{34} \\ x_{15} & x_{25} & x_{35} \end{bmatrix}$$

lineáris forma az

$$x_{11} + x_{12} + x_{13} + x_{14} + x_{15} = 9,$$

$$x_{21} + x_{22} + x_{23} + x_{24} + x_{25} = 4,$$

$$x_{31} + x_{32} + x_{33} + x_{34} + x_{35} = 8,$$

$$x_{11} + x_{21} + x_{31} = 3$$

$$x_{12} + x_{22} + x_{32} = 5$$

$$x_{13} + x_{23} + x_{33} = 4$$

$$x_{14} + x_{24} + x_{34} = 6$$

$$x_{15} + x_{25} + x_{35} = 3$$

$$x_{ij} \geq 0$$

feltételek mellett az

$$(23) \quad \begin{aligned} x_{11} &= 0, & x_{12} &= 0, & x_{13} &= 4, & x_{14} &= 5, & x_{15} &= 0 \\ x_{21} &= 0, & x_{22} &= 4, & x_{23} &= 0, & x_{24} &= 0, & x_{25} &= 0 \\ x_{31} &= 3, & x_{32} &= 1, & x_{33} &= 0, & x_{34} &= 1, & x_{35} &= 3 \end{aligned}$$

értékekre veszi fel minimumát és ez a minimum 150-nel egyenlő.

A hozzárendelési probléma és a szállítási probléma ily módon felismert ekvivalenciája esetleg érdekes elméletileg, de eredeti formájában semmiesetre sem ad kényelmesen használható segédeszközt a numerikus megoldásra. Avégből, hogy ezen ekvivalencia alapján egy gyakorlatilag is alkalmas módszert alakíthassunk ki, a következő megjegyzéseket bocsátjuk előre.

Az általános szállítási problémához tartozó (20) matrix az ekvivalens hozzárendelési probléma (19) matrixából származtatható azáltal, hogy bizonyos sorokat illetve oszlopokat egyetlen sorrá illetve oszloppá egyesítünk. Ezért közelfekvő az a gondolat, hogy a (20) matrix minden egyes vonalához megfelelő multiplicitást rendelünk hozzá és pedig $[c_{ij}]$ i -edik sorához a v_i multiplicitást és a j -edik oszlopához a μ_j multiplicitást rendeljük. Ennek megfelelően, ha a c_{ij} elemeknek valamely részhalmazát az i_1 -edik, i_2 -edik, \dots i_p -edik sorból és a j_1 -edik, j_2 -edik, \dots j_q -edik oszlopból álló vonalrendszer tartalmazza, akkor azt fogjuk mondani, hogy ez a részhalmaz $v_{i_1} + v_{i_2} + \dots + v_{i_p} + \mu_{j_1} + \mu_{j_2} + \dots + \mu_{j_q}$ multiplicitásösszegű vonalrendszerrel van fedve.

A c_{ij} elemeket ezúttal is iteratív lépésekben a

$$(24) \quad c_{ij}^{(\alpha+1)} = c_{ij}^{(\alpha)} - u_i^{(\alpha)} - v_j^{(\alpha)}; \quad c_{ij}^{(0)} = c_{ij}; \quad c_{ij}^{(\alpha)} \geq 0, \quad \alpha = 1, 2, \dots$$

transzformációkkal addig módosítjuk (átlagban csökkentjük), míg egy olyan transzformált matrixhoz jutunk, melynek 0-elemeit nem lehet n -nél kisebb multiplicitásösszegű vonalrendszerrel fedni.

Ezen előkészítő megjegyzések után bebizonyítjuk a következőket:

Ha a $[c_{ij}^{(\alpha)}]$ matrix összes 0-elemei fedhetők n -nél kisebb multiplicitásösszegű vonalrendszerrel, akkor megadható egy olyan (8) típusú transzformáció, mely az

$$S_\alpha = \sum_i \sum_j v_i \mu_j c_{ij}^{(\alpha)}$$

összeget legalább n -nel csökkenti.

Ily módon véges számú lépésben eljutunk egy olyan $[c_{ij}^{(\omega)}]$ matrixhoz, melynek 0-elemeit nem lehet n -nél kisebb multiplicitásösszegű vonalrendszerrel fedni. Ekkor a $[c_{ij}^{(\omega)}]$ matrix 0-elemei egyértelmű hozzárendelést létesítenek a \tilde{T}_i termelő helyek és az \tilde{F}_j fogyasztó helyek között. Továbbá ezen 0-elemek mindegyikéhez egyszerű utasítással hozzárendelhető egy olyan multiplicitás, melynél

az i -edik sorban levő 0-elemek multiplicitásösszege v_i

a j -edik oszlopban levő 0-elemek multiplicitásösszege μ_j .

Bizonyítás. Tegyük fel, hogy a $[c_{ij}^{(\alpha)}]$ matrix 0-elemei az i_1 -edik, i_2 -edik \dots i_p -edik sorból és a j_1 -edik, j_2 -edik, \dots j_q -edik oszlopból álló vonalrendszerrel fedhetők és hogy a multiplicitásösszegre fennáll

$$(25) \quad v_{i_1} + v_{i_2} + \dots + v_{i_p} + \mu_{j_1} + \mu_{j_2} + \dots + \mu_{j_q} < n$$

továbbá jelöljük $[c_{ij}^{(\alpha)}]$ legkisebb fedetlen elemét ε -nal.

Vonjunk most ki a $[c_{ij}^{(q)}]$ matrix minden eleméből ε -t és azután az i_1 -edik, i_2 -edik, \dots i_p -edik sor és a j_1 -edik, j_2 -edik, \dots j_q -edik oszlop elemeihez adjunk hozzá ε -t. (Ezek a műveletek (8) típusú transzformációk.)

Ily módon egy új $[c_{ij}^{(q+1)}]$ matrixot kapunk, melynek elemei nem-negatívak és amelyre (25) szerint

$$S_{a+1} = S_a - n \left(n - \sum v_{i_a} - \sum \mu_{j_\beta} \right) \varepsilon \leq S_a - n \varepsilon .$$

Ennek a transzformációnak az iterálása nyilván végül egy olyan $[c_{ij}^{(q)}]$ matrixot eredményez, melynek 0-elemei nem fedhetők n -nél kisebb multiplicitásösszegű vonalrendszerrel. Ez esetben azonban a fentebb idézett tétel garantálja, hogy a megfelelő hipermatrix n olyan 0-elemet tartalmaz, melyek közül nincs kettő ugyanabban a sorban vagy oszlopban. Ha mármost a $[c_{ij}^{(q)}]$ matrix mindegyik 0-eleméhez azt a multiplicitást rendeljük, amely a (19) hipermatrix megfelelő C_{ij} blokkjában levő 0-elemek számával egyenlő, akkor ezek a multiplicitások a (2), (3) feltételeknek is eleget tesznek és így a szállítási probléma megoldását szolgáltatják.

Ezen multiplicitások meghatározása tulajdonképpen ekvivalens egy új, *kompatibilis* lineáris egyenletrendszer megoldásával. Azonban — kivételes esetektől eltekintve — ezeknek a multiplicitásoknak a meghatározása könnyű feladat, mert, ha a $[c_{ij}^{(q)}]$ matrix i -edik sora (illetve j -edik oszlopa) egyetlen 0-elemet tartalmaz, akkor ezen 0-elemhez nyilván v_i (illetve μ_j) multiplicitás rendelendő. Hasonlóképp járhatunk el, ha a $[c_{ij}^{(q)}]$ matrix i -edik sora (illetve j -edik oszlopa) egyetlen olyan 0-elemet tartalmaz, melynek multiplicitása még ismeretlen, mert ekkor ehhez a 0-elemhez olyan multiplicitás rendelendő, mely a többi (már ismert) multiplicitások összegét v_i -re (illetve μ_j -re) egészíti ki.

A módszer részletes kivitelét a fentebb (21) táblázattal megadott szállítási példán mutatjuk be.

(26)

10	20	5	9	10	$v_1 = 9$
2	10	8	30	6	$v_2 = 4$
1	20	7	10	4	$v_3 = 8$
$\mu_1 = 3$	$\mu_2 = 5$	$\mu_3 = 4$	$\mu_4 = 6$	$\mu_5 = 3$	

Vonjuk ki minden sor elemeiből az illető sor legkisebb elemét :

(27)

5	15	0	4	5	9
0	8	6	28	4	4
0	19	6	9	3	8
3	5	4	6	3	

Vonjuk ki minden oszlop elemeiből az illető oszlop legkisebb elemét

$$(28) \quad [c_{ij}^{(1)}] = \begin{array}{c|c|c|c|c|c} \downarrow & 5 & 7 & 0 & 0 & \downarrow 2 & 9 \\ \hline \rightarrow & 0 & 0 & 6 & 24 & 1 & 4 \\ \hline & 0 & 11 & 6 & 5 & 0 & 8 \\ \hline & 3 & 5 & 4 & 6 & 3 & \end{array}$$

Az így nyert $[c_{ij}^{(1)}]$ matrix 0-elemei

- a 9 multiplicitású első sorral
- a 4 multiplicitású második sorral
- a 3 multiplicitású első oszloppal és
- a 3 multiplicitású ötödik oszloppal,

tehát $9 + 4 + 3 + 3 = 19$ multiplicitásösszegű vonalrendszerrel fedhetők.

Most alkalmazzuk a fentebb leírt transzformációt, tehát $[c_{ij}^{(1)}]$ összes elemeiből kivonjuk a legkisebb fedetlen elemet, azaz 5-öt és aztán az első és második sor, valamint az első és ötödik oszlop elemeihez 5-öt hozzáadunk. Ily módon adódik

$$(29) \quad [c_{ij}^{(2)}] = \begin{array}{c|c|c|c|c|c} \downarrow & 10 & 7 & 0 & 0 & \downarrow 7 & 9 \\ \hline \rightarrow & 5 & 0 & 6 & 24 & 6 & 4 \\ \hline & 0 & 6 & 1 & 0 & 0 & 8 \\ \hline & 3 & 5 & 4 & 6 & 3 & \end{array}$$

A (29) ábra szerint $[c_{ij}^{(2)}]$ 0-elemei $3 + 4 + 6 + 3 + 4 = 20$ multiplicitásösszegű vonalrendszerrel fedhetők. $[c_{ij}^{(2)}]$ összes elemeiből kivonjuk a legkisebb fedetlen elemet 6-ot és aztán az első, harmadik, negyedik, ötödik oszlop és a második sor elemeihez hozzáadunk 6-ot. Ily módon adódik

$$(30) \quad [c_{ij}^{(3)}] = \begin{array}{c|c|c|c|c|c} 10 & 1 & 0 & 0 & 7 & 9 \\ \hline 11 & 0 & 12 & 30 & 12 & 4 \\ \hline 0 & 0 & 1 & 0 & 0 & 8 \\ \hline 3 & 5 & 4 & 6 & 3 & \end{array}$$

$[c_{ij}^{(3)}]$ -elemeinek fedéséhez legalább 21 multiplicitásösszegű vonalrendszer szükséges. A fentiek szerint tehát $[c_{ij}^{(3)}]$ már a szállítási probléma megold-

dását fogja szolgáltatni. $[c_{ij}^{(9)}]$ 0-elemei a termelő helyek és a fogyasztó helyek egymáshoz rendelését határozzák meg. Ezután még az egyes 0-elemek multiplicitását kell megkeresnünk. Az első, harmadik és ötödik oszlop, valamint a második sor egyetlen 0-át tartalmaz. Tehát 0_{31} , 0_{22} , 0_{13} és 0_{35} multiplicitásai rögtön megadhatók. (Lásd a (31) ábrát, ahol a multiplicitások megfelelő számú ponttal vannak ábrázolva.)

0_{32} multiplicitása 4-et 5-re és 0_{14} multiplicitása 4-et 9-re egészíti ki. Végül a 0_{34} elemhez nyilván 1 multiplicitást kell hozzárendelni.

(31)

10	20	5	9	10	9
2	10	8	20	6	4
1	20	7	10	4	8
3	5	4	6	3	

A szállítási költségeknek beírása után a (31). ábrát nyerjük, melyből a minimális szállítási összköltségre (23)-mal megegyezésben

$$3 \cdot 1 + 4 \cdot 10 + 1 \cdot 20 + 4 \cdot 5 + 5 \cdot 9 + 1 \cdot 10 + 3 \cdot 4 = 150$$

adódik.

A gyakorlatilag fontos feladat, melynél az egyes termelő helyekről elszállítható egységek száma egyenlőtlenségek által van korlátozva, a fenti eljárás csekély módosításával megoldható, mégpedig úgy, hogy a szállítási költségek $[c_{ij}]$ mátrixához egy 0-okból álló oszlopot csatolunk. Ennek a 0 oszlopnak a beiktatása azt fejezi ki, hogy a termékek egy részével kapcsolatban szállítási költségek nem merülnek fel, mert egy része a termékeknek a termelő helyeken visszamarad.

A módszer általánosításának bemutatása céljából induljunk ki a fenti példából, melynek adatait annyiban módosítjuk, hogy a \tilde{T}_1 , \tilde{T}_2 , \tilde{T}_3 termelőhelyek mindegyikénél 9 tárolt egységet tételezünk fel. Ekkor a szállítási probléma táblázata a következő lesz

(32)

10	20	5	9	10	0	9
2	10	8	30	6	0	9
1	20	7	10	4	0	9
3	5	4	6	3	6	

(Az utolsó oszlop alján beírt 6 egység a termelő helyeken szükségképpen visszamaradó terméket jelenti.)

Ez esetben minden sor tartalmaz már egy 0-elemet, tehát a kezdeti transzformáció abban fog állni, hogy minden oszlop elemeiből levonjuk az oszlop legkisebb elemét:

$$(33) \quad [c_{ij}^{(1)}] = \begin{array}{c|c|c|c|c|c|c} \downarrow & \downarrow & & & \downarrow & \downarrow & \\ \rightarrow & 9 & 10 & 0 & 0 & 6 & 0 & 9 \\ \hline & 1 & 0 & 3 & 21 & 2 & 0 & 9 \\ \hline & 0 & 10 & 2 & 1 & 0 & 0 & 9 \\ \hline & 3 & 5 & 4 & 6 & 3 & 6 & \end{array}$$

Az így nyert $[c_{ij}^{(1)}]$ matrix 0-elemei fedhetők a (33). ábrán feltüntetett vonalrendszerrel, melynek multiplicitásösszege 26 (< 27), tehát a (8) típusú transzformáció alkalmazható. Minden elemből levonunk 1-et és azután a fedő sorok és fedő oszlopok minden eleméhez hozzáadunk 1-et. Ily módon adódik

$$[c_{ij}^{(2)}] = \begin{array}{c|c|c|c|c|c|c} & 10 & 11 & 0 & 0 & 7 & 1 & 9 \\ \hline & 1 & 0 & 2 & 20 & 2 & 0 & 9 \\ \hline & 0 & 10 & 1 & 0 & 0 & 0 & 9 \\ \hline & 3 & 5 & 4 & 6 & 3 & 6 & \end{array}$$

melynek elemei 27-nél kisebb multiplicitásösszegű vonalrendszerrel nem fedhetők. Eszerint $[c_{ij}^{(2)}]$ 0-elemei már a termelő helyek és fogyasztó helyek egymáshoz rendelését adják. Továbbá a fenti utasítással a 0_{31} , 0_{22} , 0_{13} és 0_{35} 0-elemek multiplicitása rögtön adódik, majd ezek segítségével 0_{14} és 0_{25} multiplicitása, mindezek alapján 0_{36} multiplicitása, végül 0_{34} multiplicitása.

A szállítási költség beírásával ezúttal az alábbi

$$\begin{array}{c|c|c|c|c|c|c} & 10 & 20 & 5 & 9 & 10 & 0 & 9 \\ \hline & 2 & 10 & 8 & 30 & 6 & 0 & 9 \\ \hline & 1 & 20 & 7 & 10 & 4 & 0 & 9 \\ \hline & 3 & 5 & 4 & 6 & 3 & 6 & \end{array}$$

táblázatot nyerjük, mely szerint a minimális szállítási költség most 140-nel egyenlő.

Nyilván az összköltség csökkenése ez esetben annak a következménye, hogy a termékek szállítása szempontjából a raktárkészletekre vonatkozó egyenlőtlenségek kevesebb megkötést jelentenek, mint az előbbi példában szereplő egyenletek.

Ha végül a fenti példa adatait azzal a módosítással tartjuk meg, hogy a \tilde{T}_1 , \tilde{T}_2 , \tilde{T}_3 termelő helyek mindegyike 10 (vagy több) egységet tárol, akkor

az optimális szállítási költség tovább csökken. Ennek az esetnek ugyanis az alábbi táblázat felel meg

10	20	5	9	10	0	10
2	10	8	30	6	0	10
1	20	7	10	4	0	10
3	5	4	6	3	9	

(A termelési helyeken 9 egység marad vissza.) Ezúttal is minden oszlopból levonva az illető oszlop legkisebb elemét, adódik

9	10	0	0	6	0	10
1	0	3	21	2	0	10
0	10	2	1	0	0	10
3	5	4	6	3	9	

Az így nyert matrix 0-elemeinek fedéséhez legalább 30 multiplicitás-összegű vonalrendszer szükséges. Ez a matrix tehát már megadja a keresett hozzárendelést és miután az első öt oszlop mindegyike egyetlen 0-elemet tartalmaz, ezeknek a multiplicitásai azonnal felírhatók és azt a triviális tényt szemléltetik, hogy mindegyik fogyasztó hely egész szükségletét arról a termelő helyről kell ellátni, ahonnan a szállítási költség minimális. Végül az alábbi táblázat mutatja,

10	20	5	9	10	0	10
2	10	8	30	6	0	10
1	20	7	10	4	0	10
3	5	4	6	3	9	

hogy a minimális szállítási összköltség ez esetben 139.

(Beérkezett : 1958. november 18.)

IRODALOM

- [1] EGERVÁRY, J. : „Matrixok kombinatorikus tulajdonságairól.” *Matematikai és Fizikai Lapok* **38** (1931) 16–28. Lásd még : KÜHN, H. W.—TUCKER, A. W. (editors) : *Linear inequalities*. Princeton University Press, Princeton (New Jersey), 1956. Bibliography, [92].

- [2] KUHN, H. W.: „On the problem of shortest path between points.” *Bulletin of the American Mathematical Society* 59 (1953).
 [3] FLOOD, M. M.: „The travelling-salesman problem.” *Journal of the Operations Research* 4 (1956).
 [4] EGERVÁRY, E.: „Bemerkung zum Transportproblem.” *Mathematik, Technik und Wirtschaft-Mitteilungen* (1958) 278—284.

КОМБИНАТОРНЫЙ МЕТОД ДЛЯ РЕШЕНИЯ ПРОБЛЕМЫ ПЕРЕВОЗКИ

J. EGERVÁRY

Резюме

Из матрицы n -ого порядка $[a_{ij}]$ состоящей из неотрицательных целых элементов, можно составить $n!$ различных сумм вида $a_{1v_1} + a_{2v_2} + \dots + a_{nv_n}$. В проблеме сопоставления (assignment problem) нужно определить наименьшую из этих сумм. Решение тривиально, если матрица содержит n таких 0-элементов, никакая пара которых не принадлежит одному и тому же столбцу или строке. В настоящей работе автор — используя опубликованную им в 1931-ом году [1] комбинаторную теорему — сводит общий случай к этому очевидному случаю.

Автор доказывает далее, что общая проблема перевозки (предполагая что все фигурирующие в ней данные — целые числа) может быть решена аналогичным методом.

Метод, опубликованный в работе, согласно свидетельству опытов, требует значительно меньше вычислительной работы, чем другие известные методы (например, метод симплексов).

COMBINATORICAL METHOD FOR SOLVING THE TRANSPORT PROBLEM

by

E. EGERVÁRY

Abstract

Let $[a_{ij}]$ be a matrix of order n whose elements are non-negative integers. One can form $n!$ different „permutation sums” $a_{1v_1} + a_{2v_2} + \dots + a_{nv_n}$ (where v_1, v_2, \dots, v_n is a permutation of the numbers $1, 2, \dots, n$), and the so called assignment problems consists in determining the least such sum. The solution is trivial if the matrix contains n elements which are equal to 0 and are such that no two of them are lying in the same row or column. The general case can be reduced to this trivial case by using a combinatorial theorem of the author, published in 1931 (see [1]).

In the present paper it is shown that the general transport problem can also be solved by a similar method (provided that all data are nonnegative integers) by attributing multiplicities to the rows and columns of the matrix of the transport problem.

Experience has shown that the method published in the present paper requires less numerical work than other known methods (e. g. the simplex method).

ON SOME JOINT DISTRIBUTIONS AND JOINT LIMITING DISTRIBUTIONS IN THE THEORY OF ORDER STATISTICS, II.

by

ISTVÁN VINCZE

Introduction

Let $\xi_1, \xi_2, \dots, \xi_n$ and $\eta_1, \eta_2, \dots, \eta_n$ be samples taken from populations having the distribution functions $F(x)$ and $G(x)$ resp. Let us denote further by $F_n(x)$ resp. $G_n(x)$ the empirical distribution functions of the samples. As usual put

$$D_{n,n}^+ = \max_{-\infty < x < \infty} (F_n(x) - G_n(x)) \quad \text{and} \quad D_{n,n} = \max_{-\infty < x < \infty} |F_n(x) - G_n(x)|.$$

Let us denote in the following by $\xi_0^{(n)}$, resp. $\eta_0^{(n)}$ the „first” points, where the functions $F_n(x) - G_n(x)$, resp. $|F_n(x) - G_n(x)|$ take their maximum, that is the values $D_{n,n}^+$, resp. $D_{n,n}$. If $\zeta_1^* \leq \zeta_2^* \leq \dots \leq \zeta_{2n}^*$ represent the union of the above two samples rearranged according to their order of magnitude, further

$$R_{n,n}^+ = \frac{1}{2} (F_n(\xi_0^{(n)} + 0) + G_n(\xi_0^{(n)} + 0))$$

and

$$R_{n,n} = \frac{1}{2} (F_n(\eta_0^{(n)} + 0) + G_n(\eta_0^{(n)} + 0)),$$

then $2n R_{n,n}^+$, resp. $2n R_{n,n}$ are the first indices in the series ζ_i^* , where the deviations take their maximum values $D_{n,n}^+$ and $D_{n,n}$ resp.

In a recent paper [3] the joint distribution and joint limiting distribution of the pairs of random variables $(D_{n,n}^+, R_{n,n}^+)$ and $(D_{n,n}, R_{n,n})$, further the limiting distribution of the random variable

$$\sqrt{\frac{n}{2}} \frac{D_{n,n}^+}{R_{n,n}^+}$$

was determined under the assumption $F(x) \equiv G(x)$. The present paper contains some further results regarding the one sided finite case.

§ 1. Summary of the results

1. According to Theorem 1. of our mentioned paper, in case $F(x) \equiv G(x)$

$$P_{r,k}^{(n)} = \mathbf{P} \left\{ D_{n,n}^+ = \frac{k}{n}, R_{n,n}^+ = \frac{r}{2n} \right\} =$$

$$= \begin{cases} 0, & \text{if } k < 0 \text{ or } k+r \text{ odd} \\ \frac{1}{(r-1)(2n-r+2)} \frac{\binom{r}{2} \binom{2n-r}{n-\frac{r}{2}}}{\binom{2n}{n}}, & \text{if } k=0, r=2, 4, \dots, 2n, \\ \frac{k(k+1)}{r(2n-r+1)} \frac{\binom{r}{2} \binom{2n-r+1}{n-\frac{r+k}{2}}}{\binom{2n}{n}}, & \text{if } k=1, 2, \dots, n; r=k, \\ & k+2, \dots, 2n-k. \end{cases}$$

In case $k=0$ the first maximum place would be $\xi_0^{(n)} = -\infty$ or the lower limit of the variable resp.; in this case in determining $\xi_0^{(n)}$ we do not regard the first interval, for which $F_n(x) - G_n(x) = 0$, but only those x values which are larger than a point x_0 , for which $F_n(x_0) - G_n(x_0) < 0$.

In § 2. the generating function of the probabilities $P_{r,k}^{(n)}$ is determined. We obtain in case $k=0$ the formula

$$(1.1) \quad \sum_{n=1}^{\infty} \sum_{(r)}^* \binom{2n}{n} P_{r,k}^{(n)} v^r w^n = \frac{4v^2 w}{(1 + \sqrt{1-4v^2 w})(1 + \sqrt{1-4w})},$$

where $\sum_{(r)}^*$ denotes the summation for $r = 2, 4, \dots, 2n$. For $k \geq 1$ the following formula holds:

$$(1.2) \quad \sum_{n=k}^{\infty} \sum_{(r)}^* \binom{2n}{n} P_{r,k}^{(n)} v^r w^n = \frac{2 \cdot 4^k \cdot v^k w^k}{(1 + \sqrt{1-4v^2 w})^k (1 + \sqrt{1-4w})^{k+1}},$$

where $\sum_{(r)}^*$ denotes now summation for $r = k, k+2, \dots, 2n-k$. (1.1) and (1.2) give the generating function of the three variables r, k, n . Using the notation $P_{0,0}^{(n)} = 0$, we obtain

$$\sum_{n=0}^{\infty} \sum_{k=0}^n \sum_{(r)}^* \binom{2n}{n} P_{r,k}^{(n)} u^k v^r w^n =$$

$$= \frac{4vw}{1 + \sqrt{1-4w}} \left\{ \frac{v}{1 + \sqrt{1-4v^2 w}} + \frac{2u}{(1 + \sqrt{1-4w})(1 + \sqrt{1-4v^2 w}) - 4uvw} \right\}.$$

In determining the generating functions we make use of the following relations, proved in a very simple, elementary way: for $k = 0, 1, 2, \dots$

$$\sum_{s=0}^{\infty} \binom{k+2s}{s} v^s = \frac{1}{\sqrt{1-4v}} \left(\frac{2}{1+\sqrt{1-4v}} \right)^k,$$

$$\sum_{s=k}^{\infty} \binom{-k+2s}{s} v^s = \frac{1}{\sqrt{1-4v}} \left(\frac{2v}{1+\sqrt{1-4v}} \right)^k.$$

In the last point of § 2. we give a table for the probabilities $P_{r,k}^{(n)}$ in case $n = 10$.

2. The marginal distribution of the probabilities $P_{r,k}^{(n)}$ regarding k gives the distribution of GNEDENKO—KOROLJUK [1]. As a consequence we obtained in our mentioned paper the probabilistic proof of the following identities

$$(1.3) \quad \sum_{(r)}^* P_{r,k}^{(n)} = \frac{2k+1}{n+k+1} \frac{\binom{2n}{n-k}}{\binom{2n}{n}},$$

where the summation is to be extended for $r = k, k+2, \dots, 2n-k$. In § 3. we give a direct proof of the identities (1.3) by means of the generating functions.

In the same §. the conditional expectation of the random variable $R_{n,n}^+$, under the condition $D_{n,n}^+ = \frac{k}{n}$ is determined:

$$\mathbf{M} \left\{ R_{n,n}^+ | D_{n,n}^+ = \frac{k}{n} \right\} = \begin{cases} \frac{1}{2} \left(1 + \frac{1}{n} \right), & \text{if } k=0, \\ \frac{k}{2k+1} \left(1 + \frac{1}{2n} \right), & \text{if } k=1, 2, \dots, n. \end{cases}$$

3. Let us denote by $P_{r,k}^{(n)}$ the probability that the maximum value $D_{n,n}^+$ happens for the first time at the term ζ_r^* , i. e. r is the "order" of the maximum deviation:

$$\sum_{(k)}^* P_{r,k}^{(n)} = P_{r,\cdot}^{(n)}.$$

We prove in §. 4. the following interesting relation valid for $q = 1, 2, \dots$

$$(1.4) \quad P_{2q-1,\cdot}^{(n)} = P_{2q,\cdot}^{(n)}.$$

We shall show, that this relations means the validity of the identities :
if r is even ($r = 2, 4, \dots, 2n$), then

$$\begin{aligned}
 (1.4') \quad & \frac{1}{(r-1)(2n-r+1)} \binom{r}{\frac{r}{2}} \binom{2n-r}{n-\frac{r}{2}} + \\
 & + \sum_{l=1}^{\min(\frac{r}{2}, n-\frac{r}{2})} \frac{4l(2l+1)}{r(2n-r+2l+2)} \binom{r}{\frac{r}{2}+l} \binom{2n-r}{n-\frac{r}{2}-l} = \\
 & = \sum_{l=1}^{\min(\frac{r}{2}, n-\frac{r}{2})} \frac{4l(2l-1)}{(r-1)(2n-r+2l+2)} \binom{r-1}{\frac{r}{2}+l-1} \binom{2n-r+1}{n-\frac{r}{2}-l+1}.
 \end{aligned}$$

Relation (1.4) is proved with the aid of the generating function of the probabilities $P_{r,n}^{(n)}$ which has the form¹

$$\sum_{n=1}^{\infty} \sum_{r=1}^{2n} \binom{2n}{n} P_{r,n}^{(n)} v^r w^n = \frac{v}{1-v} \frac{\sqrt{1-4v^2w} - \sqrt{1-4w}}{1 + \sqrt{1-4w}}.$$

§ 2. Determination of the generating functions in the one sided case

1. We have to determine the generating function

$$G(u, v, w) = \sum_{(n)} \sum_{(k)} \sum_{(r)} \binom{2n}{n} P_{r,k}^{(n)} u^k v^r w^n,$$

where

$$(2.1) \quad P_{r,0}^{(n)} = \frac{1}{(r-1)(2n-r+2)} \frac{\binom{r}{\frac{r}{2}} \binom{2n-r}{n-\frac{r}{2}}}{\binom{2n}{n}}, \quad r = 2, 4, \dots, 2n,$$

$$(2.2) \quad P_{r,k}^{(n)} = \frac{k(k+1)}{r(2n-r+1)} \frac{\binom{r}{\frac{r}{2}+k} \binom{2n-r+1}{n-\frac{r}{2}-k}}{\binom{2n}{n}}, \quad \begin{matrix} k = 1, 2, \dots, n; r = k, \\ k+2, \dots, 2n-k \end{matrix}$$

in a closed form.

¹ K. SARKADI gave a direct proof for the relations (1.4) with the aid of the random walk model, which will be published in this periodical.

We introduce instead of r the new variable $s = \frac{r-k}{2}$, where $s = 0, 1, 2, \dots, n-k$ and obtain for (2.1) and (2.2) the following forms (the latter with a slight modification)

$$(2.1') \quad P_{s,0}^{(n)} = \frac{1}{2(2s-1)(n-s+1)} \frac{\binom{2s}{s} \binom{2n-2s}{n-s}}{\binom{2n}{n}}, \quad s = 1, 2, \dots, n$$

$$(2.2') \quad P_{s,k}^{(n)} = \frac{k(k+1)}{(k+2s)(n-s+1)} \frac{\binom{k+2s}{s} \binom{-k+2(n-s)}{n-s}}{\binom{2n}{n}}, \quad s = 0, 1, 2, \dots, n-k,$$

2. At first we evaluate the following generating functions

$$g_a(v) = \sum_{s=0}^{\infty} \binom{a+2s}{s} v^s, \quad a = \dots, -2, -1, 0, 1, 2, \dots$$

and

$$G(u, v) = \sum_{a=-\infty}^{\infty} u^a \sum_{s=0}^{\infty} \binom{a+2s}{s} v^s = \sum_{a=-\infty}^{\infty} g_a(v) u^a.$$

Collecting those terms for which $a + 2s = n$ holds, we obtain:

$$G(u, v) = \sum_{n=0}^{\infty} \left(\sum_{s=0}^n \binom{n}{s} v^s u^{n-s} \right) = \sum_{n=0}^{\infty} u^{-n} (v + u^2)^n = \frac{1}{1 - \left(u + \frac{v}{u}\right)}.$$

From the relation

$$\frac{1}{1 - \left(u + \frac{v}{u}\right)} = \frac{1}{\sqrt{1-4v}} \left(\frac{1}{\frac{1 + \sqrt{1-4v}}{2} - u} + \frac{1}{u - \frac{1 - \sqrt{1-4v}}{2}} \right)$$

we get for complex values of u and v and for

$$(2.3) \quad \left| \frac{1 - \sqrt{1-4v}}{2} \right| < |u| < \left| \frac{1 + \sqrt{1-4v}}{2} \right|$$

the identity

$$\frac{1}{1 - \left(u + \frac{v}{u}\right)} = \frac{1}{\sqrt{1-4v}} \left(\sum_{a=1}^{\infty} \frac{(2u)^a}{(1 + \sqrt{1-4v})^a} \right) + \sum_{a=0}^{\infty} \frac{1}{u^a} \left(\frac{2v}{1 + \sqrt{1-4v}} \right)^a$$

So

$$(2.4) \quad \sum_{s=0}^{\infty} \binom{a+2s}{s} v^s = \frac{1}{\sqrt{1-4v}} \left(\frac{2}{1+\sqrt{1-4v}} \right)^a \text{ for } a=0, 1, 2, \dots$$

$$\sum_{s=0}^{\infty} \binom{-a+2s}{s} v^s = \frac{1}{\sqrt{1-4v}} \left(\frac{2v}{1+\sqrt{1-4v}} \right)^a \text{ for } a=0, 1, 2, \dots$$

is obtained.

For $|v| < \frac{1}{4}$ (2.3) determines an annulus for u . The row (2.4) is divergent

for $|v| \geq \frac{1}{4}$.

For $a=0$ we get from both formulae the well-known relation

$$\sum_{s=0}^{\infty} \binom{2s}{s} v^s = \frac{1}{\sqrt{1-4v}}.$$

3. In the following we prove for every integer $k \geq 1$ the relation

$$(2.5) \quad \sum_{n=k}^{\infty} \sum_{s=0}^{n-k} \binom{k+2s}{s} \binom{-k+2(n-s)}{n-s} v^s v'^{n-s} =$$

$$= \frac{1}{\sqrt{1-4v}(1+\sqrt{1-4v})^k} \frac{4^k v'^k}{\sqrt{1-4v'}(1+\sqrt{1-4v'})^k}.$$

Let us for this reason regard the product $g_k(v)g_{-k}(v')$ which can be written according to (2.4) in the form (2.5) if we collect on the left side the terms for which $s+l=n$, $l=n-s$.

4. Here we determine the generating function of the probabilities $P_{s,k}^{(n)}$ according to s and n . We write in (2.5) v^2 instead of v and multiply each side with v^{k-1} . So we get

$$(2.6) \quad \sum_{n=k}^{\infty} \sum_{s=0}^{n-k} \binom{k+2s}{s} \binom{-k+2(n-s)}{n-s} v^{k+2s-1} v'^{n-s} =$$

$$= \frac{v^{k-1}}{\sqrt{1-4v^2}(1+\sqrt{1-4v^2})^k} \frac{4^k v'^k}{\sqrt{1-4v'}(1+\sqrt{1-4v'})^k}.$$

Now we integrate with respect to v from 0 to v and with respect to v' from 0 to v' . The integration of the right side may be carried out in the following way: put

$$I_1 = \int_0^v \frac{t^{k-1}}{(1+\sqrt{1-4t^2})^{k-1}} \frac{dt}{\sqrt{1-4t^2}(1+\sqrt{1-4t^2})}.$$

Using the transformation $t' = \frac{t}{1 + \sqrt{1 - 4t^2}}$, we get

$$dt' = \frac{dt}{\sqrt{1 - 4t^2}(1 + \sqrt{1 - 4t^2})}.$$

The new integration limits are $\left(0, \frac{v}{1 + \sqrt{1 - 4v^2}}\right)$, i. e.

$$I_1 = \frac{1}{k} \left(\frac{v}{1 + \sqrt{1 - 4v^2}} \right)^k.$$

Let us now put

$$I_2 = 2 \int_0^{v'} \frac{t^k}{(1 + \sqrt{1 - 4t})^k} \frac{dt}{2\sqrt{1 - 4t}}.$$

The transformation $t' = \frac{t}{1 + \sqrt{1 - 4t}}$ leads to

$$dt' = \frac{dt}{2\sqrt{1 - 4t}}$$

and to the integration limits $\left(0, \frac{v'}{1 + \sqrt{1 - 4v'}}\right)$.

Hence

$$I_2 = \frac{2}{k+1} \left(\frac{v'}{1 + \sqrt{1 - 4v'}} \right)^{k+1}.$$

So after a slight modification we obtain from (2.6) the following identity :

$$\begin{aligned} \sum_{n=k}^{\infty} \sum_{s=0}^{n-k} \frac{k(k+1)}{(k+2s)(n-s+1)} \binom{k+2s}{s} \binom{-k+2(n-s)}{n-s} v^{2s} v'^{n-s} = \\ = \frac{2}{(1 + \sqrt{1 - 4v^2})^k} \frac{4^k v'^k}{(1 + \sqrt{1 - 4v'})^k}. \end{aligned}$$

Substituting vw instead of v^2 and w instead of v' , we obtain the generating function for $k = 1, 2, \dots$

$$(2.7) \quad \sum_{n=k}^{\infty} \sum_{s=0}^{n-k} \binom{2n}{n} P_{s,k}^{(n)} v^s w^n = \frac{2}{(1 + \sqrt{1 - 4vw})^k} \frac{4^k w^k}{(1 + \sqrt{1 - 4w})^{k+1}}.$$

5. Let us return to the case $k = 0$. As it is known

$$g_0(v) = \sum_0^{\infty} \binom{2s}{s} v^s = \frac{1}{\sqrt{1 - 4v}}.$$

Then the following relation holds

$$g_0(v) g_0(v') - g_0(v') = \sum_{n=1}^{\infty} \sum_{s=1}^n \binom{2s}{s} \binom{2n-2s}{n-s} v^s v'^{n-s}.$$

We substitute v^2 instead of v , divide by v^2 and integrate with respect to both variables. After multiplication with $\frac{v}{v'}$, we replace v^2 by v and obtain

$$\begin{aligned} \sum_{n=1}^{\infty} \sum_{s=1}^n \frac{1}{2(2s-1)(n-s+1)} \binom{2s}{s} \binom{2n-2s}{n-s} v^s v'^{n-s} = \\ = \frac{4v}{(1+\sqrt{1-4v})(1+\sqrt{1-4v'})}. \end{aligned}$$

Substituting vw instead of v and w instead of v' , we obtain the generating function for the case $k=0$:

$$(2.8) \quad \sum_{n=1}^{\infty} \sum_{s=1}^n \binom{2n}{n} P_{s,0}^{(n)} v^s w^n = \frac{4vw}{(1+\sqrt{1-4vw})(1+\sqrt{1-4w})}.$$

6. Now the generating functions for the three quantities k, s, n , resp. k, r, n may be determined.

Multiplication with u^k in (2.7), summarising for $k=1, 2, \dots$ and the addition of (2.8) results in

$$\begin{aligned} \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \sum_{s=0}^{n-k} \binom{2n}{n} P_{s,k}^{(n)} u^k v^s w^n = \\ = \frac{4w}{1+\sqrt{1-4w}} \left\{ \frac{v}{1+\sqrt{1-4vw}} + \frac{2u}{(1+\sqrt{1-4vw})(1+\sqrt{1-4w})-4uvw} \right\} \end{aligned}$$

In order to receive the generating function for k, r, n we first substitute $\frac{r-k}{2}$ instead of s , then v^2 instead of v and at last uv instead of u . So we obtain $-P_{0,0}^{(n)} = 0$

$$\begin{aligned} G(u, v, w) = \sum_{n=0}^{\infty} \sum_{k=0}^n \sum_{(r)}^* \binom{2n}{n} P_{r,k}^{(n)} u^k v^r w^n = \\ = \frac{4vw}{1+\sqrt{1-4w}} \left\{ \frac{v}{1+\sqrt{1-4v^2w}} + \frac{2u}{(1+\sqrt{1-4u^2v})(1+\sqrt{1-4w})-4uvw} \right\}. \end{aligned}$$

We wish to remark that by substituting in the formulae (2.7) and (2.8) $\frac{r-k}{2}$ instead of s , v^2 instead of v , we obtain formulas (1.1) and (1.2) of §. 1.1.

7. The following table contains the probabilities $P_{r,k}^{(n)}$ for all possible values of the pair (k, r) .

Table of the probabilities $P_{r,k}^{(10)}$

										$s=0$										
10.									$5,4 \cdot 10^{-4}$	$s=1$										
9.								$5,41 \cdot 10^{-5}$	$4,87 \cdot 10^{-5}$	$s=2$										
8.							0,0003		0,0004		0,0002	$s=3$								
7.						0,0011		0,0017		0,0015		0,0008	$s=4$							
6.					0,0034		0,0050		0,0051		0,0042		0,0023	$s=5$						
5.				0,0089		0,0116		0,0119		0,0110		0,0089		0,0054	$s=6$					
4.			0,0197		0,0217		0,0208		0,0195		0,0179		0,0155		0,0108	$s=7$				
3.		0,0383		0,0325		0,0279		0,0250		0,0235		0,0226		0,0216		0,0185	$s=8$			
2.	0,0646		0,0372		0,0272		0,0225		0,0204		0,0200		0,0209		0,0232		0,0263	$s=9$		
1.	0,0909		0,0263		0,0155		0,0116		0,0100		0,0095		0,0100		0,0116		0,0155		0,0263	$s=10$
0.		0,0263		0,0077		0,0046		0,0036		0,0032		0,0032		0,0036		0,0046		0,0077		0,0263
$\frac{k}{r}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

§. 3. Decomposition of the probabilities of Gnedenko—Koroljuk. Conditional expectations

In order to prove the identities (1.3) mentioned in § 1.2. we make use of formula (2.7) of § 2.4. which is valid for $k \geq 1$:

$$(2.7) \quad \sum_{n=k}^{\infty} \sum_{s=0}^{n-k} \binom{2n}{n} P_{s,k}^{(n)} v^s w^n = \frac{2}{(1 + \sqrt{1 - 4vw})^k} \frac{4^k w^k}{(1 + \sqrt{1 - 4w})^{k+1}}.$$

For the value $v = 1$ the coefficient of w^n in the power series will be

$$\binom{2n}{n} \sum_{s=0}^{n-k} P_{s,k}^{(n)}.$$

We make use of the following known expansion valid for every real value of a in the interval $-\frac{1}{4} < x < \frac{1}{4}$. (See e. g. [2] p. 35.)

$$(3.1) \quad (1 + \sqrt{1 + 4x})^a = 2^a \left\{ 1 + \frac{a}{1!} x + \frac{a(a-3)}{2!} x^2 + \right. \\ \left. + \frac{a(a-4)(a-5)}{3!} x^3 + \dots + \frac{a(a-l-1)(a-l-2) \dots (a-2l+1)}{a!} x^l + \dots \right\}.$$

In our case we have

$$2 \cdot 4^k w^k (1 + \sqrt{1 - 4w})^{-2k-1}.$$

So we carry out the substitutions $a = -2k - 1$, $x = -w$ and determine the coefficient of the term with the index $l = n - k$:

$$\frac{2 \cdot 4^k 2^{-2k-1} (-1)^{n-k} (-2k-1)(-2k-1-n+k+1) \dots (-2k-1-2n+2k+1)}{(n-k)!} = \\ = \frac{2k+1}{n+k+1} \binom{2n}{n-k}.$$

As this holds for $k = 1, 2, \dots$, the identity (1.3) of § 1.2. is proved in these cases.

2. In case $k = 0$ we have to verify the relation

$$\sum_{s=1}^n \frac{1}{2(2s-1)(n-s+1)} \binom{2s}{s} \binom{2n-2s}{n-s} = \frac{1}{n+1} \binom{2n}{n}.$$

This happens in a similar way as in 1. of this § with the aid of the generating function given in (2.8).

3. Now we determine the expected value of $2n R_{n,n}^+$ under the condition $D_{n,n}^+ = \frac{k}{n}$. As per definition

$$(3.2) \quad \mathbf{M} \left(2n R_{n,n}^+ \mid D_{n,n}^+ = \frac{k}{n} \right) = \frac{\sum_{r=k}^{2n-k} * r P_{r,k}^{(n)}}{\sum_{r=k}^{2n-k} * P_{r,k}^{(n)}} = \frac{\sum_{s=0}^{n-k} (k+2s) P_{s,k}^{(n)}}{\sum_{s=0}^{n-k} P_{s,k}^{(n)}} = k + 2 \frac{\sum_{s=0}^{n-k} s P_{s,k}^{(n)}}{\sum_{s=0}^{n-k} P_{s,k}^{(n)}}.$$

We evaluate the quantity $\sum_{s=0}^{n-k} s P_{s,k}^{(n)}$ for the case $k \geq 1$ with the aid of formula (2.7). After derivation with respect to w we obtain for the value $w = 1$

$$\sum_{n=k}^{\infty} \binom{2n}{n} w^n \sum_{s=0}^{n-k} s P_{s,k}^{(n)} = \frac{4^{k+1} k w^{k+1}}{\sqrt{1-4w} (1 + \sqrt{1-4w})^{2k+2}}.$$

The direct expansion of the right side did not result in a simple form. But by integrating we succeeded in obtaining a suitable expression. We carry out the integration after multiplication by w^{k+1} and get

$$\sum_{n=k}^{\infty} \frac{1}{n+k+2} \binom{2n}{n} w^{n+k+2} \sum_{s=0}^{n-k} s P_{s,k}^{(n)} = \frac{2 \cdot 4^{k+1} \cdot k}{2k+3} \frac{w^{2k+3}}{(1 + \sqrt{1-4w})^{2k+3}}.$$

Now we apply the expansion under (3.1) of this § for the case of $a = -2k-3$ and $x = -w$ and we evaluate the coefficient of the term with index $l = n - k - 1$. This will be

$$\frac{2^{-2k-3} (2k+3)}{2n+1} \binom{2n+1}{n-k-1}.$$

So we obtain the following equality

$$\frac{1}{n+k+2} \binom{2n}{n} \sum_{s=0}^{n-k} s P_{s,k}^{(n)} = \frac{k}{2n+1} \binom{2n+1}{n-k-1}$$

from which we get

$$\sum_{s=0}^{n-k} s P_{s,k}^{(n)} = k \frac{n-k}{n+k+1} \frac{\binom{2n}{n-k}}{\binom{2n}{n}}.$$

According to formula (3.2) we obtain

$$\mathbf{M} \left(2n R_{n,n}^+ \mid D_{n,n}^+ = \frac{k}{n} \right) = k + 2k \frac{k}{2k+1} = \frac{k}{2k+1} (2n+1).$$

and at last we result in

$$\mathbf{M} \left(R_{n,n}^+ | D_{n,n}^+ = \frac{k}{n} \right) = \frac{k}{2k+1} \left(1 + \frac{1}{2n} \right).$$

In the case $n \rightarrow \infty$, $k \rightarrow \infty$ it holds

$$\lim \mathbf{M} \left(R_{n,n}^+ | D_{n,n}^+ = \frac{k}{n} \right) = \frac{1}{2}.$$

4. In the case $k = 0$ a similar method leads to the result.

$$\mathbf{M}(R_{n,n}^+ | D_{n,n}^+ = 0) = \frac{1}{2} \left(1 + \frac{1}{n} \right).$$

§ 4. A property of the order of the one sided maximum deviation

1. In the following the proof is given for the relations (1.4) mentioned in § 1.3. The function $G(u, v, w)$ at the value $u = 1$ generates the probabilities $P_{r,n}^{(n)}$:

$$G(1, v, w) = \frac{4w}{1 + \sqrt{1-4w}} \left\{ \frac{v}{1 + \sqrt{1-4v^2w}} + \frac{2}{(1 + \sqrt{1-4v^2w})(1 + \sqrt{1-4w}) - 4vw} \right\}.$$

Multiplying numerator and denominator of the second term in the bracket with

$$1 - 4vw + \sqrt{1-4v^2w} - \sqrt{1-4w} - \sqrt{1-4v^2w} \sqrt{1-4w},$$

we get the following expression :

$$\frac{2(1 - 4vw + \sqrt{1-4v^2w} - \sqrt{1-4w} - \sqrt{1-4v^2w} \sqrt{1-4w})}{8w(1-v)(1 + \sqrt{1-4v^2w})}.$$

Replacing that in the formula of $G(1, v, w)$ we obtain after a slight modification

$$G(1, v, w) = \frac{v}{1-v} \frac{\sqrt{1-4v^2w} - \sqrt{1-4w}}{1 + \sqrt{1-4w}}.$$

Let us write this expression in the following forms

$$\begin{aligned} G(1, v, w) &= \frac{\sqrt{1-4v^2w} - \sqrt{1-4w}}{1 + \sqrt{1-4w}} (v + v^3 + v^5 + \dots) + \\ &+ v \frac{\sqrt{1-4v^2w} - \sqrt{1-4w}}{1 + \sqrt{1-4w}} (v + v^3 + v^5 + \dots) = \\ &= \sum_{(r)} a_{r,n} v^r w^n + \sum_{(r)} a_{r,n} v^{r+1} w^n. \end{aligned}$$

We may observe that in the second sum of this last part of our equality the exponent $(r+1)$ of v takes only even values and so in the first sum r is odd.

Hence the coefficient $a_{r,n}$ must agree with $\binom{2n}{n} P_{r,n}^{(n)}$, as well as with $\binom{2n}{n} P_{r+1,n}^{(n)}$,

which gives the proof of our statement.

2. We obtain the form (1.4') of § 1. 3. in the following way. In case of an even r , k is even too, i. e. $k = 2l$. — For an even r formula (2.1) gives the probability $P_{r,0}^{(n)}$ and the formula (2.2) gives the values of $P_{r,k}^{(n)}$ for the even $k \geq 2$. In the latter case we substituted $2l$ for k . As for k the relation $2 \leq k \leq \min(r, 2n-r)$ holds for l the inequality $1 \leq l \leq \min\left(\frac{r}{2}, n - \frac{r}{2}\right)$ is valid. For $r-1$ the term $P_{r-1,0}^{(n)}$ does not occur, and for k we may make the substitution $k = 2l-1$, $l = 1, 2, \dots, \min\left(\frac{r}{2}, n - \frac{r}{2}\right)$.

(Received 27 November, 1958)

REFERENCES

- [1] ГНЕДЕНКО, Б. В.—КОРОЛЮК, В. С.: «О максимальном расхождении двух эмпирических распределений.» Доклады Академии Наук СССР **80** (1951) 525—528.
- [2] РЫЖИК, И. М.—ГРАДШТЕЙН: Таблицы интегралов, сумм, рядов и произведений. Государственное издательство, Москва—Ленинград, 1951.
- [3] VINCZE. I.: „Einige zweidimensionale Verteilungs- und Grenzverteilungssätze in der Theorie der geordneten Stichproben.“ A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei **2** (1957) 183—209.

KÉTDIMENZIÓS ELOSZLÁS ÉS HATÁRELOSZLÁSTÉTELEK A RENDEZETT MINTÁK ELMÉLETÉBEN, II.

VINCZE I.

Kivonat

Legyenek $\xi_1, \xi_2, \dots, \xi_n$ és $\eta_1, \eta_2, \dots, \eta_n$ az $F(x)$, illetve $G(x)$ eloszlásfüggvényekkel bíró ξ , illetve η valószínűségi változókból vett n elemű minták. Jelölje $F_n(x)$ és $G_n(x)$ a megfelelő empirikus eloszlásfüggvényeket. Mint szokásos, legyenek

$$D_{n,n}^+ = \max_{-\infty < x < \infty} (F_n(x) - G_n(x))$$

és

$$D_{n,n} = \max_{-\infty < x < \infty} |F_n(x) - G_n(x)|.$$

Jelölje $\xi_0^{(n)}$, illetve $\eta_0^{(n)}$ az első pontot, ahol az $F_n(x) - G_n(x)$, illetve az $|F_n(x) - G_n(x)|$ függvények maximumukat — azaz a D_{nn}^+ , illetve D_{nn} értékeket — felveszik. Ha most $\zeta_1^*, \zeta_2^*, \dots, \zeta_{2n}^*$ a fenti két minta rendezett egyesítése, továbbá

$$R_{nn}^+ = \frac{1}{2} (F_n(\xi_0^{(n)} + 0) + G_n(\xi_0^{(n)} + 0))$$

és

$$R_{nn} = \frac{1}{2} (F_n(\eta_0^{(n)} + 0) + G_n(\eta_0^{(n)} + 0)),$$

akkor a $2n R_{nn}^+$, illetve $2n R_{nn}$ a ζ_i^* sorozatban annak az első elemnek az indexe, amelynél a D_{nn}^+ , illetve D_{nn} értékeket a megfelelő eltérések felveszik.

Szerző korábbi dolgozatában [3] meghatározta a (D_{nn}^+, R_{nn}^+) és a (D_{nn}, R_{nn}) valószínűségi változó párok együttes eloszlását és határeloszlását az $F(x) \equiv G(x)$ feltevés mellett.

Jelöljük most $P_{r,k}^{(n)}$ -rel annak valószínűségét, hogy az (egyoldali) eltérés maximuma éppen $\frac{k}{n}$ és ez először a ζ_r^* mintaelemnél következik be, vagyis legyen

$$P_{r,k}^{(n)} = \mathbf{P} \left\{ D_{nn}^+ = \frac{k}{n}, R_{nn}^+ = \frac{r}{2n} \right\}.$$

Itt $k = 0, 1, 2, \dots, n$, $r = k, k+2, \dots, 2n-k$. Megjegyezzük azonban, hogy a $k=0$ esetben a $\xi_0^{(n)} = -\infty$, vagyis $r=1$ volna, ez esetben azonban $\xi_0^{(n)}$ értelmezésénél csupán azokat az x helyeket tekintjük, amelyek nagyobbak mint valamely x_0 pont, amelyre $F(x_0) - G(x_0) < 0$. Így a $k=0$ esetben $r = 2, 4, \dots, 2n$. Vezessük még be a $P_{00}^{(n)} = 0$ jelölést is.

Dolgozatunk a $P_{r,k}^{(n)}$ valószínűségek generátorfüggvényének meghatározásával és néhány ebből levezethető eredménnyel foglalkozik. Bebizonyítjuk a következő formulákat: ha $k=0$

$$(1.1) \quad \sum_{n=1}^{\infty} \sum_{(r)}^* \binom{2n}{n} P_{r,0}^{(n)} v^r w^n = \frac{4v^2 w}{(1 + \sqrt{1-4w})(1 + \sqrt{1-4v^2 w})},$$

ahol a $\sum_{(r)}^*$ az $r = 2, 4, \dots, 2n$ -re vonatkozó összegezést jelenti; ha $k = 1, 2, \dots$, akkor

$$(1.2) \quad \sum_{n=k}^{\infty} \sum_{(r)}^* P_{r,k}^{(n)} v^r w^n = \frac{2 \cdot 4^k v^k w^k}{(1 + \sqrt{1-4w})^{k+1} (1 + \sqrt{1-4v^2 w})^k},$$

ahol most a $\sum_{(r)}^*$ az $r = k, k+2, \dots, 2n-k$ -ra vonatkozó összegezést jelenti.

(1.1) és (1.2)-ből kapjuk a három változós generátorfüggvényt:

$$\begin{aligned} & \sum_{n=0}^{\infty} \sum_{k=0}^n \sum_{r=k}^{2n-k} \binom{2n}{n} P_{r,k}^{(n)} u^k v^r w^n = \\ & = \frac{4vw}{1 + \sqrt{1-4w}} \left\{ \frac{v}{1 + \sqrt{1-4v^2 w}} + \frac{2u}{(1 + \sqrt{1-4w})(1 + \sqrt{1-4v^2 w}) - 4uvw} \right\}. \end{aligned}$$

A generátorfüggvények meghatározásához igen egyszerű módon igazoljuk az alábbi relációkat:

$$\sum_{s=0}^{\infty} \binom{k+2s}{s} v^s = \frac{1}{\sqrt{1-4v}} \left(\frac{2}{1+\sqrt{1-4v}} \right)^k,$$

$$\sum_{s=k}^{\infty} \binom{-k+2s}{s} v^s = \frac{1}{\sqrt{1-4v}} \left(\frac{2v}{1+\sqrt{1-4v}} \right)^k,$$

mindkét esetben $k = 0, 1, 2, \dots$ lehet.

A dolgozat $n = 10$ esetre a $P_{k,r}^{(n)}$ valószínűségek táblázatát adja.

A $P_{k,r}^{(n)}$ valószínűségek k -ra vonatkozó peremeloszlása a GNEDENKO—KOROLJUK-féle eloszlása [1]. Ennélfogva már előző dolgozatunkban valószínűségi számítási bizonyítását nyertük a következő azonosságoknak

$$(1.3) \quad \sum_{(r)}^* P_{r,k}^{(n)} = \frac{2k+1}{n+k+1} \frac{\binom{2n}{n-k}}{\binom{2n}{n}}, \quad k = 0, 1, 2, \dots, n,$$

míg az összegezés $r = k, k+2, \dots, 2n-k$ -ra végzendő. Ezt az azonosságot a generátorfüggvény segítségével formálisan is igazoljuk.

Meghatározzuk továbbá az $R_{n,n}^+$ valószínűségi változó feltételes várható értékét a $D_{n,n}^+ = \frac{k}{n}$ feltétel mellett. Erre nézve a következőt kapjuk:

$$\mathbf{M} \left(R_{n,n}^+ \mid D_{n,n}^+ = \frac{k}{n} \right) = \begin{cases} \frac{1}{2} \left(1 + \frac{1}{n} \right), & \text{ha } k = 0 \\ \frac{k}{2k+1} \left(1 + \frac{1}{2n} \right), & \text{ha } k = 1, 2, \dots \end{cases}$$

Jelöljük $P_{r,k}^{(n)}$ -el annak a valószínűségét, hogy a $D_{n,n}^+$ maximum éppen a ζ_r^* mintaelemnél következik be, vagyis legyen

$$\sum_{(k)}^* P_{r,k}^{(n)} = P_{r,\cdot}^{(n)}.$$

Dolgozatunkban a generátorfüggvény segítségével a következő — $q = 1, 2, \dots, n$ -re érvényes — érdekes relációt bizonyítjuk be:¹

$$(1.4) \quad P_{2q-1,\cdot}^{(n)} = P_{2q,\cdot}^{(n)}.$$

¹ Az (1.4) relációt Sarkadi Károly a bolyongási modell segítségével közvetlenül is bebizonyította; közlésére e folyóiratban kerül sor.

Ez a következő azonosságok fennállását jelenti: ha r páros szám, akkor

$$\begin{aligned}
 (1.4') \quad & \frac{1}{(r-1)(2n-r+1)} \binom{r}{\frac{r}{2}} \binom{2n-r}{n-\frac{r}{2}} + \\
 & + \sum_{l=1}^{\min(\frac{r}{2}, n-\frac{r}{2})} \frac{4l(2l+1)}{r(2n-r+2l+2)} \binom{r}{\frac{r}{2}+l} \binom{2n-r}{n-\frac{r}{2}+l} = \\
 & = \sum_{l=1}^{\min(\frac{r}{2}, n-\frac{r}{2})} \frac{4l(2l-1)}{(r-1)(n-\frac{r}{2}+2l+2)} \binom{r-1}{\frac{r}{2}+l-1} \binom{2n-r+1}{n-\frac{r}{2}-l+1}.
 \end{aligned}$$

A $P_{r,}^{(n)}$ valószínűségek generátorfüggvényére a következő formula érvényes:

$$\sum_{n=1}^{\infty} \sum_{r=1}^{2n} \binom{2n}{n} P_{r,}^{(n)} v^r w^n = \frac{v}{1-v} \frac{\sqrt{1-4v^2w} - \sqrt{1-4w}}{1 + \sqrt{1-4w}}.$$

ДВУМЕРНЫЕ ТЕОРЕМЫ О РАСПРЕДЕЛЕНИИ И ПРЕДЕЛЬНОМ РАСПРЕДЕЛЕНИИ В ТЕОРИИ ВАРИАЦИОННЫХ РЯДОВ, II.

I. VINCZE

Резюме

Пусть $\xi_1, \xi_2, \dots, \xi_n$ и $\eta_1, \eta_2, \dots, \eta_n$ суть n -элементные выборки значений случайных величин ξ и η с функциями распределения $F(x)$ и $G(x)$. Пусть $F_n(x)$ и $G_n(x)$ обозначают соответствующие эмпирические функции распределения. Как обычно, пусть

$$D_{nn}^+ = \max_{-\infty < x < \infty} (F_n(x) - G_n(x))$$

и

$$D_{nn} = \max_{-\infty < x < \infty} |F_n(x) - G_n(x)|$$

Обозначим через $\xi_0^{(n)}$ и $\eta_0^{(n)}$ первые точки, в которых $F_n(x) - G_n(x)$ и $|F_n(x) - G_n(x)|$ принимают свои наибольшие значения, т. е. D_{nn}^+ и D_{nn} . Если

теперь $\zeta_1^*, \zeta_2^*, \dots, \zeta_{2n}^*$ упорядоченное объединение вышеуказанных выборок далее

$$R_{nn}^+ = \frac{1}{2} (F_n(\xi_0^{(n)} + 0) + G_n(\xi_0^{(n)} + 0))$$

и

$$R_{nn} = \frac{1}{2} (F_n(\eta_0^{(n)} + 0) + G_n(\eta_0^{(n)} + 0)) ,$$

то $2n R_{nn}^+$ и $2n R_{nn}$ суть индексы тех первых элементов в последовательности ζ_i^* на которых соответствующие отклонения принимают значения D_{nn}^+ и D_{nn} .

Автор в предыдущей работе [3] определил совместное распределение и предельное распределение пар случайных величин (D_{nn}^+, R_{nn}^+) и (D_{nn}, R_{nn}) при условии $F(x) \equiv G(x)$.

Обозначим через $P_{k,r}^{(n)}$ вероятность того, что максимум (одностороннего) отклонения равен k/n и это имеет место впервые для элементы выборки ζ_r^* т. е. пусть

$$P_{r,k}^{(n)} = \mathbf{P} \left\{ D_{nn}^+ = \frac{k}{n}, R_{nn}^+ = \frac{r}{2n} \right\} .$$

Здесь $k = 0, 1, 2, \dots, n$; $r = k, k+2, \dots, 2n-k$. Однако, заметим, что в случае $k=0$ было бы $\xi_0^{(n)} = -\infty$, т. е. $r=1$, но в этом случае при определении $\xi_0^{(n)}$ принимаются во внимание лишь те x , которые превосходят некоторое x_0 , для которого $F(x_0) - G(x_0) < 0$. Поэтому, в случае $k=0$, $r=2, 4, \dots, 2n$. Введём ещё обозначение $P_{00}^{(n)} = 0$.

Работа занимается определением производящей функции вероятностей $P_{kr}^{(n)}$ и некоторыми результатами, которые могут быть отсюда выведены. Доказываются следующие формулы: если $k=0$, то

$$(1.1) \quad \sum_{n=1}^{\infty} \sum_{(r)}^* \binom{2n}{n} P_{r,0}^{(n)} v^r w^n = \frac{4 v^2 w}{(1 + \sqrt{1-4w})(1 + \sqrt{1-4v^2w})}$$

где $\sum_{(r)}^*$ обозначает суммирование относительно $r = 2, 4, \dots, 2n$; если $k = 1, 2, \dots$, то

$$(1.2) \quad \sum_{n=k}^{\infty} \sum_{(r)}^* P_{r,k}^{(n)} v^r w^n = \frac{2 \cdot 4^k v^k w^k}{(1 + \sqrt{1-4w})^{k+1} (1 + \sqrt{1-4v^2w})^k} ,$$

где теперь $\sum_{(r)}^*$ обозначает суммирование относительно $r = k, k+2, \dots, 2n-k$. Из (1.1) и (1.2) получается производящая функция от трех переменных:

$$\begin{aligned} & \sum_{n=0}^{\infty} \sum_{r=0}^n \sum_{k=r}^{2n-k} \binom{2n}{n} P_{r,k}^{(n)} u^k v^r w^n = \\ & = \frac{4 v w}{1 + \sqrt{1-4w}} \left\{ \frac{v}{1 + \sqrt{1-4v^2w}} + \frac{2 u}{(1 + \sqrt{1-4w})(1 + \sqrt{1-4v^2w}) - 4 u v w} \right\} . \end{aligned}$$

Для определения производящих функций очень простым способом доказываются следующие соотношения:

$$\sum_{s=0}^{\infty} \binom{k+2s}{s} v^s = \frac{1}{\sqrt{1-4v}} \left(\frac{2}{1+\sqrt{1-4v}} \right)^k,$$

$$\sum_{s=k}^{\infty} \binom{-k+2s}{s} v^s = \frac{1}{\sqrt{1-4v}} \left(\frac{2v}{1+\sqrt{1-4v}} \right)^k,$$

в обоих случаях может быть $k = 0, 1, 2, \dots$

Работа дает таблицу вероятностей $P_{r,k}^{(n)}$ для случая $n = 10$.

Проекция распределении вероятностей $P_{kr}^{(n)}$ относительно k есть распределение Гнеденко—Королюка [1]. Таким образом мы уже в предыдущей работе пришли к теоретико-вероятностному доказательству следующих тождеств:

$$(1.3) \quad \sum_{(r)}^* P_{rk}^{(n)} = \frac{2k+1}{n+k+1} \frac{\binom{2n}{n-k}}{\binom{2n}{n}}, \quad k = 0, 1, 2, \dots, n$$

где суммирование производится для $r = k, k+2, \dots, 2n-k$. Это тождество с помощью производящей функции доказывается и формально.

Определяется также условное математическое ожидание случайной величины R_{nn}^+ при условии $D_{nn}^+ = k/n$. Результат таков:

$$\mathbf{M} \left\{ R_{nn}^+ \mid D_{nn}^+ = \frac{k}{n} \right\} = \begin{cases} \frac{1}{2} \left(1 + \frac{1}{n} \right), & \text{если } k = 0 \\ \frac{k}{2k+1} \left(1 + \frac{1}{2n} \right), & \text{если } k = 1, 2, \dots \end{cases}$$

Обозначим через $P_{r,\cdot}^{(n)}$ вероятность того, что максимум D_{nn}^+ имеет место как раз для элемента образца ζ_r^* т. е. пусть

$$\sum_{(k)}^* P_{rk}^{(n)} = P_{r,\cdot}^{(n)}$$

В работе с помощью производящей функции для $q = 1, 2, \dots, n$ — доказывается следующее интересное соотношение:¹

$$(1.4) \quad P_{2q-1,\cdot}^{(n)} = P_{2q,\cdot}^{(n)}$$

¹ К. SARKADI доказал соотношение (1.4) и непосредственно с помощью модели блуждания, доказательство будет опубликовано в настоящем журнале.

Это означает выполнение следующих соотношений: если r четное число, то

$$\begin{aligned}
 (1.4) \quad & \frac{1}{(r-1)(2n-r+1)} \binom{r}{2} \binom{2n-r}{n-\frac{r}{2}} + \\
 & + \sum_{l=1}^{\min(\frac{r}{2}, n-\frac{r}{2})} \frac{4l(2l+1)}{r(2n-r+2l+2)} \binom{r}{\frac{r}{2}+l} \binom{2n-r}{n-\frac{r}{2}+l} = \\
 & = \sum_{l=1}^{\min(\frac{r}{2}, n-\frac{r}{2})} \frac{4l(2l-1)}{(r-1)(n-\frac{r}{2}+2l+2)} \binom{r-1}{\frac{r}{2}+l-1} \binom{2n-r+1}{n-\frac{r}{2}-l+1}.
 \end{aligned}$$

Для производящей функции вероятностей $P_{r,\cdot}^{(n)}$ имеет место следующая формула:

$$\sum_{n=1}^{\infty} \sum_{r=1}^{2n} \binom{2n}{n} P_{r,\cdot}^{(n)} v^r w^n = \frac{v}{1-v} \frac{\sqrt{1-4v^2w} - \sqrt{1-4w}}{1 + \sqrt{1-4w}}.$$

ON THE CENTRAL LIMIT THEOREM FOR SAMPLES FROM A FINITE POPULATION

by

PAUL ERDŐS and ALFRÉD RÉNYI

Let a_1, a_2, \dots, a_n be arbitrary real numbers. Let us consider all possible $\binom{n}{s}$ sums

$$a_{i_1} + a_{i_2} + \dots + a_{i_s}, 1 \leq i_1 < i_2 < \dots < i_s \leq n$$

formed by choosing s arbitrary different elements of the sequence a_1, a_2, \dots, a_n . Let us put

$$M_n = \sum_{k=1}^n a_k$$

and

$$D_n = \sqrt{\sum_{k=1}^n \left(a_k - \frac{M_n}{n} \right)^2},$$

further

$$D_{n,s} = \sqrt{\frac{s}{n} \left(1 - \frac{s}{n} \right)} \cdot D_n.$$

Let $N_{n,s}(x)$ denote the number of those sums $a_{i_1} + a_{i_2} + \dots + a_{i_s}$ which do not exceed

$$\frac{sM_n}{n} + x D_{n,s}$$

and put

$$F_{n,s}(x) = \frac{N_{n,s}(x)}{\binom{n}{s}}.$$

We ask about conditions concerning the sequence $\{a_1, \dots, a_n\}$ and s under which $F_{n,s}(x)$ will be approximately equal to the normal distribution function $\Phi(x)$ defined by

$$(1) \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

In what follows we shall suppose throughout that

$$M_n = 0.$$

This is no essential restriction, because if $M_n \neq 0$ we may consider instead of the sequence $\{a_1, \dots, a_n\}$ the sequence $\{a'_1, \dots, a'_n\}$ where $a'_k = a_k - M_n/n$, and for the new sequence a'_k we have evidently

$$M'_n = \sum_{k=1}^n a'_k = 0.$$

We may also suppose $s \leq n/2$, because evidently if $M_n = 0$,

$$\sum_{i=1}^s a_{i_k} = - \sum_{l=1}^{n-s} a_{j_l}$$

where j_1, j_2, \dots, j_{n-s} are those among the numbers $1, 2, \dots, n$ which are not contained in the sequence i_1, i_2, \dots, i_s and therefore

$$F_{n,n-s}(x) = 1 - F_{n,s}(-x).$$

The problem can be interpreted statistically as follows: We choose a sample $\{a_{i_1}, a_{i_2}, \dots, a_{i_s}\}$ of size s from the finite population $\{a_1, a_2, \dots, a_n\}$ and ask under which conditions will the mean

$$\frac{1}{s} \sum_{j=1}^s a_{i_j}$$

of the sample be distributed asymptotically normally about the mean $\frac{M_n}{n}$ of the population. We shall prove that $F_{n,s}(x)$ tends to $\Phi(x)$, e. g. when $n \rightarrow +\infty$ and $s = s_n \rightarrow +\infty$ in such a manner that putting

$$L_n = \frac{\left(\frac{1}{n} \sum_{k=1}^n |a_k|^3 \right)^2}{\left(\frac{1}{n} \sum_{k=1}^n a_k^2 \right)^3}$$

we have

$$\lim_{n \rightarrow +\infty} \frac{L_n}{s_n} = 0;$$

or more generally, if putting

$$d_{n,s}(\varepsilon) = \frac{1}{D_n^2} \sum_{\substack{1 \leq k \leq n \\ |a_k| > \varepsilon D_n s}} a_k^2$$

we have for any $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} d_{n,s_n}(\varepsilon) = 0.$$

Clearly our question asks about the distribution of a sum of weakly dependent random variables. As a matter of fact if we imagine the elements of our sample as chosen successively without replacement and ξ_k is the element of the sequence which has been chosen at the k -th choice, then we have to investigate the distribution of

$$\frac{1}{s} \sum_{k=1}^s \xi_k,$$

the random variables ξ_1, \dots, ξ_s being not independent. (They are however equivalent.)

The special case, when all a_k are equal to α or $\beta \neq \alpha$, the number of those a_k which are equal to α being a positive percentage < 1 of n , is well-known; as a matter of fact, in this case our question reduces to the investigation of the limiting distribution of the hypergeometric distribution, and it is known (see e. g. BERNSTEIN [1]) that the Moivre—Laplace theorem can be extended for the hypergeometric (instead of the binomial) distribution.

We shall express our principal result (Theorem 1.) as relating to the n th row of an infinite triangular matrix with each row-sum being equal to 0; we deduce from this theorem statements concerning sampling among the first n terms of an infinite sequence (Theorem 2.).

We shall prove first the following.

Theorem 1. *Let us consider an infinite triangular matrix*

$$\begin{array}{ccccccc} & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ a_{n1} & a_{n2} & . & . & . & . & a_{nn} \\ & . & . & . & . & . & . \end{array}$$

with real elements, such that the sum of elements in each row is zero;

$$(2) \quad \sum_{k=1}^n a_{nk} = 0.$$

Let us put

$$(3) \quad D_n = \sqrt{\sum_{k=1}^n a_{nk}^2}$$

and

$$(4) \quad D_{n,s} = \sqrt{\frac{s}{n} \left(1 - \frac{s}{n}\right)} D_n, \quad 1 \leq s \leq n$$

where s is an integer.

Let $N_{n,s}(x)$ for any real x denote the number of those sums

$$a_{ni_1} + a_{ni_2} + \dots + a_{ni_s} \quad 1 \leq i_1 < i_2 < \dots < i_s \leq n$$

the value of which does not exceed $x D_{n,s}$, i. e. put

$$(5) \quad N_{n,s}(x) = \sum_{\substack{a_{ni_1} + \dots + a_{ni_s} < x D_{n,s} \\ 1 \leq i_1 < i_2 < \dots < i_s \leq n}} 1$$

and put

$$(6) \quad F_{n,s}(x) = \frac{N_{n,s}(x)}{\binom{n}{s}}$$

Let us put for any $\varepsilon > 0$,

$$(7) \quad d_{n,s}(\varepsilon) = \frac{1}{D_n^2} \sum_{|a_{n,k}| > \varepsilon D_{n,s}} a_{nk}^2$$

If $s_n \leq n/2$ is chosen in such a manner that when n tends to infinity we have for any $\varepsilon > 0$

$$(8) \quad \lim_{n \rightarrow \infty} d_{n,s_n}(\varepsilon) = 0,$$

it follows that for any real x we have

$$(9) \quad \lim_{n \rightarrow +\infty} F_{n,s_n}(x) = \Phi(x)$$

where $\Phi(x)$ is the normal distribution function defined by (1).

Remark. Clearly if we modify our problem as follows: we denote by $N_{n,s}^*(x)$ the number of those sums $a_{ni_1} + a_{ni_2} + \dots + a_{ni_s}$ which are $< x D_{n,s}$, where $1 \leq i_k \leq n$ ($k = 1, 2, \dots, n$), then the question whether

$$\lim_{n \rightarrow \infty} \frac{N_{n,s_n}^*(x)}{n^s} = \Phi(x)$$

holds, can be answered according to the following theorem which is a special case of a well-known theorem (see [2], p. 103):

Let ξ_{nk} ($k = 1, 2, \dots, s$) be independently and identically distributed random variables with the distribution

$$\mathbf{P}\{\xi_{nk} = a_{nj}\} = \frac{1}{n} \quad (j = 1, 2, \dots, n).$$

Then the sums

$$\zeta_{n,s} = \frac{\sum_{j=1}^s \xi_{nj}}{D_{n,s}}$$

are in the limit for $n \rightarrow +\infty$ normally distributed if and only if (8) holds for every $\varepsilon > 0$. Thus our theorem shows that if we take a sample without replacement from a finite population, the distribution of the sample mean will be in the limit normal, under the same condition as that which is known for sampling with replacement.

Proof of Theorem 1. Let us put for any real t

$$(10) \quad \varphi_{n,s}(t) = \int_{-\infty}^{+\infty} e^{itx} dF_{n,s}(x) = \frac{1}{\binom{n}{s}} \sum_{1 \leq i_1 < i_2 < \dots < i_s \leq n} e^{it(a_{ni_1} + a_{ni_2} + \dots + a_{ni_s})}.$$

To prove Theorem 1. we have to show that if (8) holds, we have

$$(11) \quad \lim_{n \rightarrow +\infty} \varphi_{n,s_n} \left(\frac{t}{D_{n,s_n}} \right) = e^{-\frac{t^2}{2}}.$$

Now evidently, putting

$$(12) \quad p = \frac{s_n}{n}$$

and

$$(13) \quad B_{n,s_n}(p) = \binom{n}{s_n} p^{s_n} (1-p)^{n-s_n}$$

we have

$$(14) \quad \varphi_{n,s_n}(u) = \frac{1}{2\pi B_{n,s_n}(p)} \int_{-\pi}^{+\pi} \left[\prod_{k=1}^n (1-p + pe^{i(\theta + ua_{nk})}) \right] e^{-i\theta s_n} d\theta.$$

Taking (2) into account we can write

$$(15) \quad \varphi_{n,s_n}(u) = \frac{1}{2\pi B_{n,s_n}(p)} \int_{-\pi}^{+\pi} \prod_{k=1}^n ((1-p)e^{-ip(\theta + ua_{nk})} + pe^{i(1-p)(\theta + ua_{nk})}) d\theta.$$

Now as well-known, if $n \rightarrow +\infty$ and $s_n \rightarrow +\infty$ (which follows clearly from (8))

$$(16) \quad B_{n,s_n}(p) \sim \frac{1}{\sqrt{2\pi np(1-p)}}$$

and thus we have, substituting $u = \frac{t}{D_{n,s_n}}$ and $\theta = \frac{\psi}{\sqrt{np(1-p)}}$

$$(17) \quad \varphi_{n,s_n} \left(\frac{t}{D_{n,s_n}} \right) \sim \frac{1}{\sqrt{2\pi}} \int_{-\pi\sqrt{np(1-p)}}^{+\pi\sqrt{np(1-p)}} \left[\prod_{k=1}^n \varrho_k(\psi, t) \right] d\psi$$

where

$$(18) \quad \varrho_k(\psi, t) = (1-p)e^{-ip\left(\frac{\psi}{\sqrt{np(1-p)}} + \frac{ta_{nk}}{D_{n,s_n}}\right)} + pe^{i(1-p)\left(\frac{\psi}{\sqrt{np(1-p)}} + \frac{ta_{nk}}{D_{n,s_n}}\right)}.$$

In what follows C_1, C_2, \dots are absolute constants. Let us choose an $\varepsilon > 0$. If k is such that

$$\frac{|t| |a_{n,k}|}{D_{n,s}} < \varepsilon$$

and $|\psi| < 2\varepsilon \sqrt{np(1-p)}$, we have

$$(19) \quad \varrho_k(\psi, t) = 1 - \frac{p(1-p)}{2} \left[\frac{\psi}{\sqrt{np(1-p)}} + \frac{ta_{nk}}{D_{n,s_n}} \right]^2 (1 + \vartheta_1 \varepsilon)$$

with $|\vartheta_1| \leq C_1$. On the other hand, we have for an arbitrary k

$$(20) \quad |\varrho_k(\psi, t) - 1| \leq C_2 \left(\frac{p(1-p)\psi^2}{2n} + \frac{p(1-p)|t\psi a_{nk}|}{2\sqrt{n}D_{n,s_n}} + \frac{p(1-p)a_{nk}^2}{8D_{n,s_n}^2} \right).$$

It follows that for $|\psi| < 2\varepsilon \sqrt{np(1-p)}$ we have

$$(21) \quad \prod_{k=1}^n \varrho_k(\psi, t) = e^{-\frac{\psi^2}{2} - \frac{t^2}{2}} \cdot (1 + \eta_n)$$

where

$$|\eta_n| \leq C_3 \left(d_{n,s_n} \left(\frac{\varepsilon}{|t|} \right) + \frac{\psi^2 p(1-p) \cdot l_n}{n} \right)$$

where

$$l_n = \sum_{\substack{1 \leq k \leq n \\ |a_{nk}| > \frac{\varepsilon}{|t|} D_{n,s_n}}} 1.$$

As by the inequality of SCHWARZ we have

$$l_n \leq \frac{t^2}{\varepsilon^2 p(1-p)} d_{n,s_n} \left(\frac{\varepsilon}{|t|} \right)$$

we obtain

$$(22) \quad \lim_{n \rightarrow +\infty} \eta_n = 0.$$

On the other hand if $|\psi| \geq 2\varepsilon \sqrt{np(1-p)}$ we use the following estimates:

$$\text{if } \frac{|t| |a_{n,k}|}{D_{n,s_n}} > \varepsilon \text{ then } |\varrho_k(\psi, t)| \leq 1$$

and

$$\text{if } \frac{|a_{n,k}| |t|}{D_{n,s_n}} \leq \varepsilon, \text{ we have } |\varrho_k(\psi, t)| \leq 1 - p(1-p)(1 - \cos \varepsilon)$$

and thus

$$\left| \int_{2\varepsilon\sqrt{np(1-p)} \leq |\psi| \leq \pi\sqrt{np(1-p)}} \sum_{k=1}^n \varrho_k(\psi, t) d\psi \right| \leq 2\pi \left(1 - \frac{C(\varepsilon)s_n}{n} \right)^n \sqrt{s_n} \leq 2\pi e^{-C(\varepsilon)s_n} \sqrt{s_n}$$

where $C(\varepsilon) > 0$ is a constant depending only on ε .

Thus (11) holds and therewith our theorem is proved.

Remark. Clearly our condition (8) is a condition of Lindeberg's type. It is easy to see that the following condition of Liapounoff's type implies (8):

$$(23) \quad \lim_{n \rightarrow +\infty} \sqrt{\frac{n}{s_n}} \cdot \frac{\left(\sum_{k=1}^n |a_{nk}|^3 \right)}{\left(\sum_{k=1}^n a_{nk}^2 \right)^{3/2}} = 0.$$

As a matter of fact, if $\frac{s_n}{n} \leq \frac{1}{2}$ we have

$$d_{n,s_n}(\varepsilon) \leq \frac{1}{\varepsilon D_{n,s_n}} \frac{\sum_{k=1}^n |a_{nk}|^3}{\sum_{k=1}^n a_{nk}^2} = \frac{1}{\varepsilon} \sqrt{\frac{2n}{s_n}} \cdot \frac{\left(\sum_{k=1}^n |a_{nk}|^3 \right)}{\left(\sum_{k=1}^n a_{nk}^2 \right)^{3/2}}.$$

Thus (23) is a sufficient condition for the validity of (9).

For the special case when $a_{nk} = a_k$ does not depend on n ($k = 1, 2, \dots, n$) we obtain from Theorem 1. the following

Theorem 2. Let $\{a_n\}$ be an arbitrary sequence of real numbers. Put

$$(24) \quad M_n = \sum_{k=1}^n a_k$$

$$D_n = \sqrt{\sum_{k=1}^n \left(a_k - \frac{M_n}{n} \right)^2}$$

and

$$D_{n,s} = \sqrt{\frac{s}{n} \left(1 - \frac{s}{n} \right)} D_n.$$

Let $N_{n,s}(x)$ denote the number of those sums

$$a_{i_1} + a_{i_2} + \dots + a_{i_s} \quad 1 \leq i_1 < i_2 < \dots < i_s \leq n$$

which do not exceed

$$\frac{s M_n}{n} + x D_{n,s}$$

and put

$$(25) \quad F_{n,s}(x) = \frac{N_{n,s}(x)}{\binom{n}{s}}.$$

If putting $a'_k = a_k - \frac{M_n}{n}$, and

$$(26) \quad d_{n,s}(\varepsilon) = \frac{1}{D_n^2} \sum_{\substack{|a'_k| > \varepsilon D_{n,s} \\ 1 \leq k \leq n}} |a'_k|^2$$

if $n \rightarrow \infty$, $s = s_n \leq \frac{n}{2}$ depends in such a manner on n that for any $\varepsilon > 0$

$$(27) \quad \lim_{n \rightarrow +\infty} d_{n,s_n}(\varepsilon) = 0$$

then we have for any real x

$$(28) \quad \lim_{n \rightarrow +\infty} F_{n,s_n}(x) = \Phi(x).$$

Let us consider some examples.

Example 1. If $0 < c \leq |a_k| \leq C$, then (27) is satisfied if $s_n \rightarrow +\infty$ arbitrarily slowly.

Example 2. (27) is satisfied, provided that $s_n \rightarrow +\infty$ arbitrary slowly if $0 < c k^\alpha \leq |a_k| \leq C k^\alpha$ with $\alpha > -1/2$.

Example 3. If

$$a_k = \frac{(-1)^k}{\sqrt{k}},$$

then (27) is satisfied if

$$\lim_{n \rightarrow +\infty} \frac{\log s_n}{\log n} = 1,$$

e. g. for

$$s_n \sim \frac{n}{(\log n)^A}$$

with arbitrary large $A > 0$.

Example 4. Evidently it is impossible to satisfy (8) if

$$\sum a_{nk}^2 \leq \alpha$$

and

$$\max_k |a_{nk}| \geq \beta > 0.$$

The most far reaching previous result known to us is that of W. G. MADOW [3]. His theorem, which can be compared with our result can be stated with our notations as follows:

If

$$(29) \quad \lambda_{nr} = \frac{\left| \frac{1}{n} \sum_{k=1}^n a_{nk}^r \right|}{\left(\frac{1}{n} \sum_{k=1}^n a_{nk}^2 \right)^{r/2}} \leq C \text{ for all } n \text{ and } r = 3, 4, \dots$$

then the assertion of Theorem 1. holds, provided that $S_n \rightarrow +\infty$. Clearly Madow's theorem follows from ours because if (29) holds for $r = 4$, we have by Hölder's inequality

$$(30) \quad \sqrt{\frac{n}{s_n} \frac{\sum_{k=1}^n |a_{nk}|^3}{\left(\sum_{k=1}^n a_{nk}^2 \right)^{3/2}}} \leq \frac{\lambda_{n4}^{3/4}}{\sqrt{s_n}} \leq \frac{C^{3/4}}{\sqrt{s_n}}$$

and thus if $s_n \rightarrow +\infty$, (23) is satisfied, which as remarked above ensures the validity of (9). Thus if (29) holds for the single value $r = 4$, our theorem can be applied, which implies that Madow's result is much weaker than ours.

Madow's result includes as a special case a previous result of F. N. DAVID [4]. Thus our theorem contains as a special case that of the paper [4] too.

(Received 1. December, 1958.)

REFERENCES

- [1] БЕРНШТЕИН, С. Н.: *Теория вероятностей*. Москва, Гостехиздат, 1945.
- [2] ГНЕДЕНКО, Б. В и КОЛМОГОРОВ, А. Н.: *Предельные распределения для сумм независимых случайных величин*. Москва, Гостехиздат, 1949.
- [3] MADOW, W. G.: „On the limiting distributions of estimates based on samples from finite universes.” *Annals of Math. Stat.* 19 (1948) 535—545.
- [4] DAVID, F. N.: „Limiting distributions connected with certain methods of sampling human population.” *Stat. Res. Mem.* (1938) pp. 69—99. especially p. 77.

VÉGES SOKASÁGBÓL VETT MINTÁKRA VONATKOZÓ CENTRÁLIS HATÁRELOSZLÁSTÉTEL

ERDŐS P. és RÉNYI A.

Kivonat

A dolgozat a valószínűségszámítás centrális határeloszlástételének bizonyos gyengén függő változók esetére való kiterjesztésével foglalkozik, amelyben kimutatja, hogy a Lindeberg-féle feltétellel analóg feltétel teljesülése esetén egy véges sokaságból vett minta elemeinek középértéke közelítőleg normális eloszlású. Más szóval a szerzők a következő tételt bizonyítják be:

1. tétel: Tekintsünk egy valós elemű

$$\begin{pmatrix} a_{11} & & & & \\ a_{21} & a_{22} & & & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ a_{n1} & a_{n2} & \cdot & \cdot & \cdot & a_{nn} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

háromszög alakú mátrixot, melyben az egy sorban álló elemek összege mindenütt nulla :

$$(2) \quad \sum_{k=1}^n a_{nk} = 0 .$$

Vezessük be a

$$(3) \quad D_n = \sqrt{\sum_{k=1}^n a_{nk}^2}$$

és

$$(4) \quad D_{n,s} = \sqrt{\frac{s}{n} \left(1 - \frac{s}{n}\right)} D_n$$

jelöléseket.

Legyen tetszőleges valós x -re $N_{n,s}(x)$ azon

$$a_{ni_1} + a_{ni_2} + \dots + a_{ni_s} \quad 1 \leq i_1 < i_2 < \dots < i_s \leq n$$

alakú összegek száma, melyek kisebbek $x D_{n,s}$ -nél, azaz

$$(5) \quad N_{n,s}(x) = \sum_{\substack{a_{ni_1} + \dots + a_{ni_s} < x D_{n,s} \\ 1 \leq i_1 < \dots < i_s \leq n}} 1$$

és legyen

$$(6) \quad F_{n,s}(x) = \frac{N_{n,s}(x)}{\binom{n}{s}} ,$$

Vezessük be a

$$(7) \quad d_{n,s}(\varepsilon) = \frac{1}{D_n^2} \sum_{|a_{nk}| > \varepsilon D_{n,s}} a_{nk}^2 \quad 1 \leq s \leq n$$

jelölést. Ha az $\{s_n\}$ ($s_n \leq n/2$) sorozatot úgy választjuk, hogy minden pozitív ε -ra

$$(8) \quad \lim_{n \rightarrow \infty} d_{n,s_n}(\varepsilon) = 0$$

legyen, akkor minden valós x -re

$$\lim_{n \rightarrow \infty} F_{n,s_n}(x) = \Phi(x) ,$$

ahol $\Phi(x)$ a normális eloszlásfüggvény.

Az 1. tétel speciális eseteként nyerhető az alábbi

2. tétel. Legyen $\{a_n\}$ tetszőleges valós számsorozat. Vezessük be az

$$(24) \quad \begin{aligned} M_n &= \sum_{k=1}^n a_k \\ D_n &= \sqrt{\sum_{k=1}^n \left(a_k - \frac{M_n}{n} \right)^2} \\ D_{n,s} &= \sqrt{\frac{s}{n} \left(1 - \frac{s}{n} \right)} D_n \end{aligned}$$

jelöléseket. Jelölje $N_{n,s}(x)$ azon

$$a_{i_1} + a_{i_2} + \dots + a_{i_s} \quad 1 \leq i_1 < i_2 < \dots < i_s \leq n$$

alakú összegek számát, melyek kisebbek, mint

$$s \frac{M_n}{n} + x D_{n,s}$$

és legyen

$$(25) \quad F_{n,s}(x) = \frac{N_{n,s}(x)}{\binom{n}{s}}.$$

Vezessük be az $a'_k = a_k - \frac{M_n}{n}$, és

$$(26) \quad d_{n,s}(\varepsilon) = \frac{1}{D_n^2} \sum_{\substack{|a'_k| > \varepsilon D_{n,s} \\ 1 \leq k \leq n}} a'_k{}^2$$

jelölést. Ha az $\{s_n\}$ ($s_n \leq n/2$) sorozatot úgy választjuk, hogy minden pozitív ε -ra

$$(27) \quad \lim_{n \rightarrow \infty} d_{n,s}(\varepsilon) = 0$$

akkor minden valós x -re

$$(28) \quad \lim_{n \rightarrow \infty} F_{n,s_n}(x) = \Phi(x).$$

ЦЕНТРАЛЬНАЯ ПРЕДЕЛЬНАЯ ТЕОРЕМА ДЛЯ ВЫБОРКАХ ВЗЯТЫХ ИЗ КОНЕЧНЫХ МНОЖЕСТВ

P. ERDŐS и A. RÉNYI

Резюме

Работа занимается распространением центральной предельной теоремы вероятностей на случай некоторых слабо зависящих случайных величин. Авторы доказывают, что в случае выполнения условия, аналогичного условию Lindeberg-a, среднее значение элементов выборки взятой из конечного множества, имеет приблизительно нормальное распределение. Иначе говоря доказывается следующая теорема:

Теорема 1: Рассмотрим треугольную матрицу

$$\begin{pmatrix} a_{11} & & & & \\ a_{21} & a_{22} & & & \\ & & \ddots & & \\ & & & \ddots & \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \\ & & & & \ddots \\ & & & & & \ddots \end{pmatrix}$$

из вещественных элементов, в которой сумма элементов каждой строки равна нулю

$$(2) \quad \sum_{k=1}^n a_{nk} = 0,$$

Введем обозначения

$$(3) \quad D_n = \sqrt{\sum_{k=1}^n a_{nk}^2}$$

и

$$(4) \quad D_{n,s} = \sqrt{\frac{s}{n} \left(1 - \frac{s}{n}\right)} D_n \quad 1 \leq s \leq n.$$

Пусть для любого вещественного x $N_{n,s}(x)$ обозначает число тех сумм вида $a_{ni_1} + a_{ni_2} + \dots + a_{ni_s}$ где $1 \leq i_1 < i_2 < \dots < i_s \leq n$, которые меньше чем $x D_{n,s}$ т. е.

$$(5) \quad N_{n,s}(x) = \sum_{\substack{a_{ni_1} + \dots + a_{ni_s} < x D_{n,s} \\ 1 \leq i_1 < i_2 < \dots < i_s \leq n}} 1$$

и пусть

$$(6) \quad F_{n,s}(x) = \frac{N_{n,s}(x)}{\binom{n}{s}}.$$

Введем обозначение

$$(7) \quad d_{n,s}(\varepsilon) = \frac{1}{D_n^2} \sum_{|a_{nk}| > \varepsilon D_{n,s}} a_{nk}^2$$

Если последовательность s_n ($\leq n/2$) выбрана так, что для каждого положительного ε выполняется условие

$$(8) \quad \lim_{n \rightarrow \infty} d_{n,s_n}(\varepsilon) = 0$$

то для всех вещественных x

$$\lim_{n \rightarrow +\infty} F_{n,s_n}(x) = \Phi(x)$$

где $\Phi(x)$ нормальная функция распределения.

В качестве специального случая теоремы 1 получается следующая

Теорема 2. Пусть $\{a_n\}$ любая последовательность вещественных чисел. Введем следующие обозначения

$$\begin{aligned} M_n &= \sum_{k=1}^n a_k \\ (24) \quad D_n &= \sqrt{\sum_{k=1}^n \left(a_k - \frac{M_n}{n} \right)^2} \\ D_{n,s} &= \sqrt{\frac{s}{n} \left(1 - \frac{s}{n} \right)} D_n \end{aligned}$$

Обозначим через $N_{n,s}(x)$ число тех сумм вида

$$a_{i_1} + a_{i_2} + \dots + a_{i_s} \quad (1 \leq i_1 < \dots < i_s \leq n)$$

которые меньше, чем

$$s \frac{M_n}{n} + x D_{n,s},$$

и пусть

$$(25) \quad F_{n,s}(x) = \frac{N_{n,s}(x)}{\binom{n}{s}}.$$

Введём обозначение $a'_k = a_k - \frac{M_n}{n}$, и

$$(26) \quad d_{n,s}(\varepsilon) = \frac{1}{D_n^2} \sum_{\substack{|a'_k| > \varepsilon D_{n,s} \\ 1 \leq k \leq n}} a'_k{}^2$$

Если последовательность s_n ($\leq n/2$) выбрана так, что для каждого положительного ε выполняется условие

$$(27) \quad \lim_{n \rightarrow \infty} d_{n,s_n}(\varepsilon) = 0$$

то для всех вещественных x

$$(28) \quad \lim_{n \rightarrow \infty} F_{n,s_n}(x) = \Phi(x).$$

THE REPRESENTATION OF A PERMUTATION AS THE PRODUCT OF A MINIMAL NUMBER OF TRANSPOSITIONS, AND ITS CONNECTION WITH THE THEORY OF GRAPHS

by

JÓZSEF DÉNES

Introduction

In the papers [5], [6] HURWITZ deals with the following problem : What is the number of representations of a given permutation of degree n as the product of w transpositions? Earlier in a special case the problem was treated by KASTEN [15].

CAYLEY [2], DZIOBEK [3], PÓLYA [10], PRÜFER [11] CLARKE [14] and others [1], [8] proved, that the problem of connected graphs with n labelled vertices¹ V_1, V_2, \dots, V_n and having $n - 1$ edges (i. e. the number of trees of order n) is n^{n-2} (see KÖNIG [7], p. 78).

At the first glance one does not see any connection between the above mentioned two problems. Nevertheless they are closely connected. The aim of the present paper is to establish the connection between the above mentioned two problems and to prove as the consequence of this connection, some further results.

In § 2 it is shown that the number of representations of a cyclical permutation of degree n as the product of $n - 1$ transpositions (this is the minimal number of transpositions required) is equal to the number of connected graphs with n given vertices and having $n - 1$ edges, and thus is equal to n^{n-2} .

§ 3 deals with some recurrence formulae concerning the numbers of representations of permutations by means of transpositions.

In § 4 it is shown, that the number of all representations of all permutations of degree n , which consist of k cycles ($1 \leq k \leq n$), as the product of a minimal number of transpositions (this number being $n - k$) is equal to the number of graphs with n given labelled vertices which consist of k disjoint trees as subgraphs.

§ 1. Definitions, notations

A) Graphs. A *graph* is defined informally as a finite collection of points (*vertices*) and a collection of *edges*, each connecting two points. By other words a graph is a set of points $S = (V_1, V_2, \dots, V_n)$ and a set of (unordered) pairs (V_i, V_j) of elements of the set S .

We restrict ourselves to consider only such graphs in which any two points are joined by at most one edge; we suppose also that the graphs has

¹ i. e. the vertices of the graphs are supposed to be distinguishable, e. g. numbered by the numbers $1, 2, \dots, n$.

no "slings", i. e. edges joining a point with itself, although in the general theory of graphs (see KÖNIG [7]), neither of these restrictions is made. A *circle* of a graph is a sequence of edges of the following form: $V_{i_1} V_{i_2}, V_{i_2} V_{i_3}, \dots, V_{i_{k-2}} V_{i_{k-1}}, V_{i_{k-1}} V_{i_1}$, where $V_i V_j$ ($i \neq j$) denotes the edge joining the points V_i and V_j , and the points $V_{i_1}, V_{i_2}, \dots, V_{i_{k-1}}$ are all different. A graph is *connected* if every pair of points is joined by a path, that is, for any i and $j \neq i$ there can be found different indices k_0, k_1, \dots, k_r such that $k_0 = i$, $k_r = j$ and all the edges $V_{k_i} V_{k_{i+1}}$ ($i = 0, 1, \dots, r-1$) belong to the graph. A connected graph without circles is called a *tree*.

B) Permutations. A *permutation* of the elements $1, 2, \dots, n$ (a permutation of degree n) can be represented by a function $f(k)$ ($k = 1, 2, \dots, n$) which takes on any of the values $1, 2, \dots, n$ exactly once. A permutation represented by the function $f(k)$ is called a *cycle* if there exists a reordering a_1, a_2, \dots, a_n of the numbers $1, 2, \dots, n$ such that $f(a_i) = a_{i+1}$ ($i = 1, 2, \dots, n$) where by definition $a_{n+1} = a_1$. The *length* of a cycle is equal to the number of the elements of the cycle. The cycle defined by $f(a_i) = a_{i+1}$ shall be denoted by $\langle a_1, a_2, \dots, a_n \rangle$. A cycle of length 2 is called a *transposition*. The *product* $P_1 P_2$ of the two permutations P_1 and P_2 represented by the functions $f_1(k)$ resp. $f_2(k)$ is defined as the permutation P_3 represented by the function $f_3(k) = f_2(f_1(k))$. The set of all permutations of the elements $1, 2, \dots, n$ forms a group, the so called *symmetric group of degree n* . If P is a permutation of a subset of the set $(1, 2, \dots, n)$ it may be extended to the whole set $(1, 2, \dots, n)$ by putting $f(i) = i$ for those i for which P is not defined. Any permutation of the elements $1, 2, \dots, n$ can be represented as the product of cycles of disjoint subsets of the elements $1, 2, \dots, n$ (SPEISER [13], p. 105), the cycles being extended to the whole set $(1, 2, \dots, n)$ as mentioned above. A permutation consists of k cycles if in its representation as a product of cycles without common element, exactly k cycles occur. The cycle of unit length is also a cycle and thus a permutation of degree n consists of n cycles if and only if it is the identical permutation. Let T denote the set of all transpositions which occur in the symmetric group of degree n . Any subset T_1 of T is called a *set of transpositions*. If T_1 is a set of transpositions, with elements t_1, t_2, \dots , it is denoted by $T_1 = (t_1, t_2, \dots)$. The *element-set* of a set of transpositions is the set of letters which occur in the transpositions. The element-set of a set $T_1 = (t_1, t_2, \dots)$ of transpositions shall be denoted by $\{T_1\} = \{t_1, t_2, \dots\}$. A set of transpositions is called *regular*² if it can not be obtained as the union of sets of transpositions with disjoint element-sets. A product of transpositions is called *regular* if the set of those transpositions which occur in the product is regular. If P is a permutation of the elements $1, 2, \dots, n$ represented by the function $f(k)$, we define the element-set of P as the set of those numbers k for which $f(k) \neq k$.

§ 2. A connection between the theory of trees and the theory of permutations

Lemma 1. A cycle of degree n can not be represented by the product of less than $n - 1$ transpositions. A product of $n - 1$ transpositions of the elements $1, 2, \dots, n$ is equal to a cycle of degree n if and only if the product is regular.

² This notion was called by HOYER and PICCARD in the papers [4] and [9] "Zusammenhängende Reihe" resp. "domaines de connexion".

Proof. First we prove that a regular product of $n - 1$ transpositions of the elements $1, 2, \dots, n$ is always equal to a cycle of degree n . Clearly the above mentioned product of $n - 1$ transpositions is a permutation, because all transpositions occurring in the product are elements of the symmetric group of degree n . If the product in question would not be a cycle then the set of transpositions occurring in this product could be splitted into subsets having disjoint element-sets, in contradiction to the definition of regularity of a product of transpositions.

Conversely the set of transpositions figuring in a representation of a cycle of degree n by means of a product of transpositions is always regular. As a matter of fact if the set of transpositions occurring in the product could be splitted into subsets having disjoint element-sets, then clearly the product of the transpositions in question would not be a cycle as it would be equal to the product of two permutations with disjoint element-sets.

It remains to prove that a cycle of degree n can not be represented by the product of less than $n - 1$ transpositions. We prove still more, namely that a permutation of degree n which has k cycles can not be represented by means of a product of less than $n - k$ transpositions.

Let us suppose in contrary to our statement

$$P_k = t_1 \cdot t_2 \dots t_{n-k-i} \quad (i \geq 1)$$

where P_k is a permutation of degree n having k cycles, and $t_1, t_2, \dots, t_{n-k-i}$ are transpositions. Now we use a theorem of SERRET [12] (p. 230) according to which if P is a permutation of degree n having k cycles and t a transposition, then the number of cycles of Pt similarly of tP is $k + 1$ or $k - 1$ according to whether the elements of t belong to the same cycle of P or not.

By this theorem $t_{n-k-i} \dots t_2 \cdot t_1 \cdot P_k = P'$ has at most $n - i$ cycles. But P' is clearly the identical permutation having exactly n cycles; thus we obtained a contradiction which proves our assertion.

Theorem 1. *The number of representations of a given cycle of degree n by means of a product of a minimal number of transpositions is equal to the number of trees with n labelled vertices.*

Proof. We establish a one-to-one correspondence between all trees with vertices V_1, V_2, \dots, V_n and all regular sets of $n - 1$ transpositions of the elements $1, 2, \dots, n$. This correspondence is obtained as follows: To any edge joining the points V_i and V_j let correspond the transposition (i, j) . Then to any tree there corresponds a regular set of $n - 1$ transpositions.³ Let us denote the number of all regular transposition-sets consisting of $n - 1$ transpositions of the elements $1, 2, \dots, n$ by T_n . T_n is clearly equal to the number of trees with labelled vertices, i. e. $T_n = n^{n-2}$. Any set of $n - 1$ transpositions may be arranged in $(n - 1)!$ distinct ways to a product.

³ This correspondence between sets of transpositions and trees has been utilized in a similar but somewhat different way, by PÓLYA in [10] pp. 208–209. where it is mentioned that to any set of $n - 1$ transpositions of the elements $1, 2, \dots, n$ which generate the whole symmetric group of degree n , (e. g. in our terminology: to any regular set of transpositions) there corresponds a tree of degree n (with *unlabelled* vertices) so that to conjugate sets of transpositions there corresponds the same tree (topologically equivalent trees being considered as identical).

Taking into consideration all these possible arrangements we obtain that the number of regular products of $n - 1$ transpositions is equal to $(n - 1)! T_n$.

Let C_n denote the number of representations of a given cycle of degree n by means of a product of $n - 1$ transpositions. As clearly the number of different cycles of degree n is equal to $(n - 1)!$, we obtain that the number of all representations of all cycles by products of $n - 1$ transpositions is equal to $(n - 1)! C_n$.

On the other hand, according to Lemma 1 any regular product of $n - 1$ transpositions is equal to a cycle of degree n and conversely, the product of $n - 1$ transpositions which is equal to a cycle of degree n is regular.

Thus we obtain that

$$(n - 1)! T_n = (n - 1)! C_n$$

and therefore

$$T_n = C_n$$

This proves Theorem 1.

Corollary. *The number of representations of a cycle of degree n by means of product of minimal number of transpositions is n^{n-2} .*

Remark. Note that we proved the equality $T_n = C_n$ not by establishing a one-to-one correspondence between all trees with n labelled vertices and all representations as a product of $n - 1$ transpositions of a *given* cycle of degree n , but by establishing a one-to-one correspondence between all possible rearrangements of the sequence of edges of all trees and the set of all representations as a product of $n - 1$ transpositions of *all* possible cycles of degree n .

It is an unsolved problem to find a direct one-to-one correspondence between the trees with n vertices and the representations of a given cycle of degree n by means of the product of $n - 1$ transpositions.

§ 3. Some recurrence formulae

Theorem 2. *Let C_n denote the number of representations of a given cycle of degree n by means of a product of $n - 1$ transpositions and $C_{(i, n-i)}$ the number of representations of a given permutation of degree n having two cycles of the length i resp. $n - i$ by means of a product of a minimal number of transpositions (*i. e.* by a product of $n - 2$ transpositions) then*

$$(1) \quad C_n = \frac{n}{2} \sum_{i=1}^{n-1} C_{(i, n-i)}.$$

Proof. Let $P_k(i, n - i)$ denote the permutation of degree n having the two cycles $\langle k + 1, \dots, k + i \rangle$ and $\langle k + i + 1, \dots, k + n \rangle$ of the length i resp. $n - i$ where $k = 0, 1, 2, \dots, n - 1$. Here $n + j$ is to be understood mod n i. e. as identical with j ($1 \leq j \leq n$).

Let us represent the permutation $P_k(i, n - i)$ in all possible ways as a product of $n - 2$ transpositions, and multiply these products from the right by the transposition $(k + 1, k + i + 1)$. We obtain in such a way a represen-

tation of the cycle $\langle 1, 2, \dots, n \rangle = \Pi$ as a product of $n - 1$ transpositions. As the two cycles of $P_k(i, n - i)$ are commutable

$$P_k(i, n - i) = P_{k+i}(n - i, i)$$

we obtain in such a way every mentioned representation twice.

Further we shall prove that in this way we obtain all representations of Π as a product of $n - 1$ transpositions. As a matter of fact if

$$\Pi = t_1 t_2 \dots t_{n-1}$$

is a representation of Π as a product of $n - 1$ transpositions and $t_{n-1} = (k + 1, k + i + 1)$, then we have $\Pi t_{n-1} = t_1 t_2 \dots t_{n-2}$ and thus $t_1 t_2 \dots t_{n-2}$ is a representation of the permutation $\Pi t_{n-1} = P_k(i, n - i)$. Thus the representation $\Pi = t_1 \dots t_{n-1}$ can be obtained by multiplying the representation $P_k(i, n - i) = t_1 t_2 \dots t_{n-2}$ by $t_{n-1} = (k + 1, k + i + 1)$. Thus we obtain

$$2C_n = n \sum_{i=1}^{n-1} C_{(i, n-i)}$$

which proves Theorem 2.

Let $C[a_1, a_2, \dots, a_n]$ where $\sum_{j=1}^n j a_j = n$ denote the number of representations of a permutation P of degree n , which has a_j cycles of length j ($j = 1, 2, \dots, n$) by means of a product of a minimal number of transpositions, i. e. a product consisting of $n - k$ transpositions where $k = \sum_{j=1}^n a_j$ is the total number of cycles of P . We now prove

Theorem 3. If $\sum_{j=1}^n j a_j = n$, $\sum_{j=1}^n a_j = k$, then

$$(2) \quad C[a_1, a_2, \dots, a_n] = (n - k)! \prod_{j=1}^n \left(\frac{C_j}{(j - 1)!} \right)^{a_j}$$

where $C_j = j^{j-2}$ is the number of representations of a cycle of length j as a product of $j - 1$ transpositions.

Proof. Let us choose for each cycle of P one of its representations as a product of a minimal number of transpositions. This can evidently be done in $\prod_{j=1}^n C_j^{a_j}$ ways. We can arrange the $n - k$ transpositions figuring in the chosen representations in such an order in which the original order of the transpositions which belong to the representation of the same cycle is preserved; in

$$\frac{(n - k)!}{\prod_{j=1}^n [(j - 1)!]^{a_j}}$$

ways. As two different transpositions are permutable if and only if they have no common element, the product of the $n - k$ transpositions in any of these arrangements is equal to P . As conversely any representation of P as a product of $n - k$ transpositions is clearly of the type just considered, Theorem 3 is proved.

Remark. As a consequence of the Corollary of Theorem 1 and Theorem 2 we can obtain the following identity :

$$(3) \quad n^{n-2} = \frac{n}{2} \sum_{j=1}^{n-1} \binom{n-2}{j-1} j^{j-2} (n-j)^{n-j-2}$$

DZIOBEK [3] (see also BOL [1]) proved a similar identity, namely

$$(4) \quad n^{n-2} = \sum_{j=1}^{n-1} \binom{n-2}{j-1} j^{j-2} (n-j)^{n-j-1}$$

by means of a formula of ABEL.

It is easy to show that the two identities are equivalent.

§ 4. Connection between representations of certain permutations by means of products of transpositions and a certain class of graphs

Theorem 4. Let $T_k(n)$ denote the number of products of $n - k$ transpositions of n elements, which are equal to some permutation consisting of k cycles, then we have

$$(5) \quad T_k(n) = n! (n - k)! \sum_{\substack{\sum_{j=1}^n a_j = k \\ \sum_{j=1}^n j a_j = n}} \prod_{j=1}^n \frac{\binom{j-2}{a_j}}{a_j!}$$

Proof. According to Theorem 3 the number of representations of a given permutation consisting of a_j cycles of length j for $j = 1, 2, \dots, n$ where $\sum_{j=1}^n j a_j = n$ and $\sum_{j=1}^n a_j = k$ is

$$(n - k)! \prod_{j=1}^n \left(\frac{C_j}{(j-1)!} \right)^{a_j}.$$

According to a well-known theorem of the theory of groups (see SPEISER [13] p. 108) the number of different permutations which have a_j cycles of length j ($j = 1, 2, \dots, n$) is equal to

$$\frac{n!}{\prod_{j=1}^n j^{a_j} a_j!}.$$

Thus we obtain

$$T_k(n) = n! (n-k)! \sum_{\substack{\sum_{j=1}^n ja_j = n \\ \sum_{j=1}^n a_j = k}} \prod_{j=1}^n \frac{\left(\frac{j^{j-2}}{(j-1)!} \right)^{a_j}}{a_j! j^{a_j}}.$$

This proves Theorem 4.

Theorem 5. If $G_k(n)$ denotes the number of graphs with n labelled vertices consisting of k disjoint trees as subgraphs, then

$$G_k(n) = n! \sum_{\substack{\sum_{j=1}^n a_j = k \\ \sum_{j=1}^n ja_j = n}} \prod_{j=1}^n \frac{\left(\frac{j^{j-2}}{j!} \right)^{a_j}}{a_j!}.$$

Proof. Evidently the number of sets of transpositions such that their product represents a permutation consisting of k cycles is equal to $T_k(n)/(n-k)!$. It is easy to show that there is a one-to-one correspondence between the sets of $n-k$ transpositions such that their product represents a permutation consisting of k cycles and between the graphs with n labelled vertices consisting of k disjoint trees.

The above mentioned correspondence is a straightforward generalization of the correspondence which we constructed in course of the proof of Theorem 1.

Thus it follows that

$$G_k(n) = \frac{T_k(n)}{(n-k)!}$$

This proves Theorem 5.

The author did not succeed to bring the expression for $G_k(n)$ to a simpler form.

(Received 16. December, 1958)

REFERENCES

- [1] BOL, G.: „Über eine kombinatorische Frage”. *Abhandlungen aus dem Mathematischen Seminar der Hansischen Universität* **12/2** (1938) 242.
- [2] CAYLEY, A.: “A theorem on trees”. *Quarterly Journal of Pure and Applied Mathematics* **23** (1889) 376—378. See also CAYLEY, A.: *Mathematical Papers* (Vol. XIII., 26—28).
- [3] DZIOBEK, O.: „Eine Formel der Substitutionstheorie”. *Sitzungsberichte der Berliner Mathematischen Gesellschaft* **17** (1917) 64—67.
- [4] HOYER, P.: “Über den Zusammenhang in Reihen mit einer Anwendung auf die Theorie der Substitutionen”. *Mathematische Annalen* **42** (1893) 58—88.
- [5] HURWITZ, A.: „Über Riemannsche Flächen mit gegebenen Verzweigungspunkte”. *Mathematische Annalen* **39** (1891) 1—61. See also HURWITZ, A.: *Mathematische Werke*. Birkhäuser, Basel, 1932. (Bd. I., 321—383).
- [6] HURWITZ, A.: “Über die Anzahl der Riemannschen Flächen mit gegebenen Verzweigungspunkte”. *Mathematische Annalen* **55** (1902) 53—66. See also HURWITZ, A.: *Mathematische Werke*. Birkhäuser, Basel, 1932, I. 492—505.
- [7] KÖNIG, D.: *Theorie der endlichen und unendlichen Graphen*. Akademische Verlag, Leipzig, 1936. Reprinted: Chelsea, New-York, 1950.
- [8] NEVILLE, E. H.: “The codifying of tree-structure”. *Proceedings of the Cambridge Philosophical Society* **49** (1953) 381—385.
- [9] PICCARD, S.: *Sur les bases des groupes d'ordre fini*. Neuchatel, 1957.
- [10] PÓLYA, G.: “Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen”. *Acta Mathematica (Uppsala)* **68** (1937) 145—255.
- [11] PRÜFER, H.: “Neuer Beweis eines Satzes über Permutationen”. *Archiv der Mathematik und Physik* **27/3** (1918) 142—144.
- [12] SERRET, J. H.: *Handbuch der höheren Algebra*. Teubner, Leipzig, 1879 (2. Auflage).
- [13] SPEISER, A.: *Die Theorie der Gruppen von endlicher Ordnung*. Birkhäuser, Basel, 1956.
- [14] CLARKE, L. E.: „On Cayley's formula for counting trees”. *Journal of the London Mathematical Society* **33** (1958) 471—474.
- [15] KASTEN, H.: *Zur Theorie der dreiblättrigen Riemannschen Flächen*. Diss. Göttingen, 1876.

PERMUTÁCIÓK MINIMÁLIS SZÁMÚ TRANSZPOZÍCIÓ SZORZATÁRA BONTÁSAINAK SZÁMA ÉS EZEK KAPCSOLATA A GRÁFELMÉLETTEL

DÉNES J.

Kivonat

CAYLEY [3] cikkében található tétel szerint az n szögpontú fák száma n^{n-2} .

Ebben a cikkben a szerző megmutatja az: n szögpontú fák száma megegyezik egy adott n -edfokú ciklus $n - 1$ transzpozíció szorzatára bontásainak számával (1. tétel).

Legyen C_n egy adott n -edfokú ciklus $n - 1$ transzpozíció szorzatára bontásainak száma, akkor $C_n = n^{n-2}$ (Korollárium).

A 2. tétel C_n -re vonatkozó rekurzív összefüggést ad.

A 4. tétel az összes k ciklust tartalmazó n -edfokú permutációk $n - k$ transzpozíció szorzatára bontásainak számát állapítja meg.

A 4. tétel gráfelméleti alkalmazása lehetővé teszi a k , páronként különböző és közös pont nélküli fák mint részgráfok álló n szögpontú gráfok számának meghatározását (5. tétel).

ЧИСЛО РАЗЛОЖЕНИЙ ПЕРЕСТАНОВОК В ПРОИЗВЕДЕНИЕ НАИ- МЕНЬШЕГО ЧИСЛА ТРАНСПОЗИЦИЙ И ИХ СВЯЗЬ С ТЕОРИЕЙ ГРАФОВ

J. DÉNES

Резюме

Согласно теореме статьи [3] САУЛЕУ-а число деревьев с n угловыми точками равно n^{n-2} .

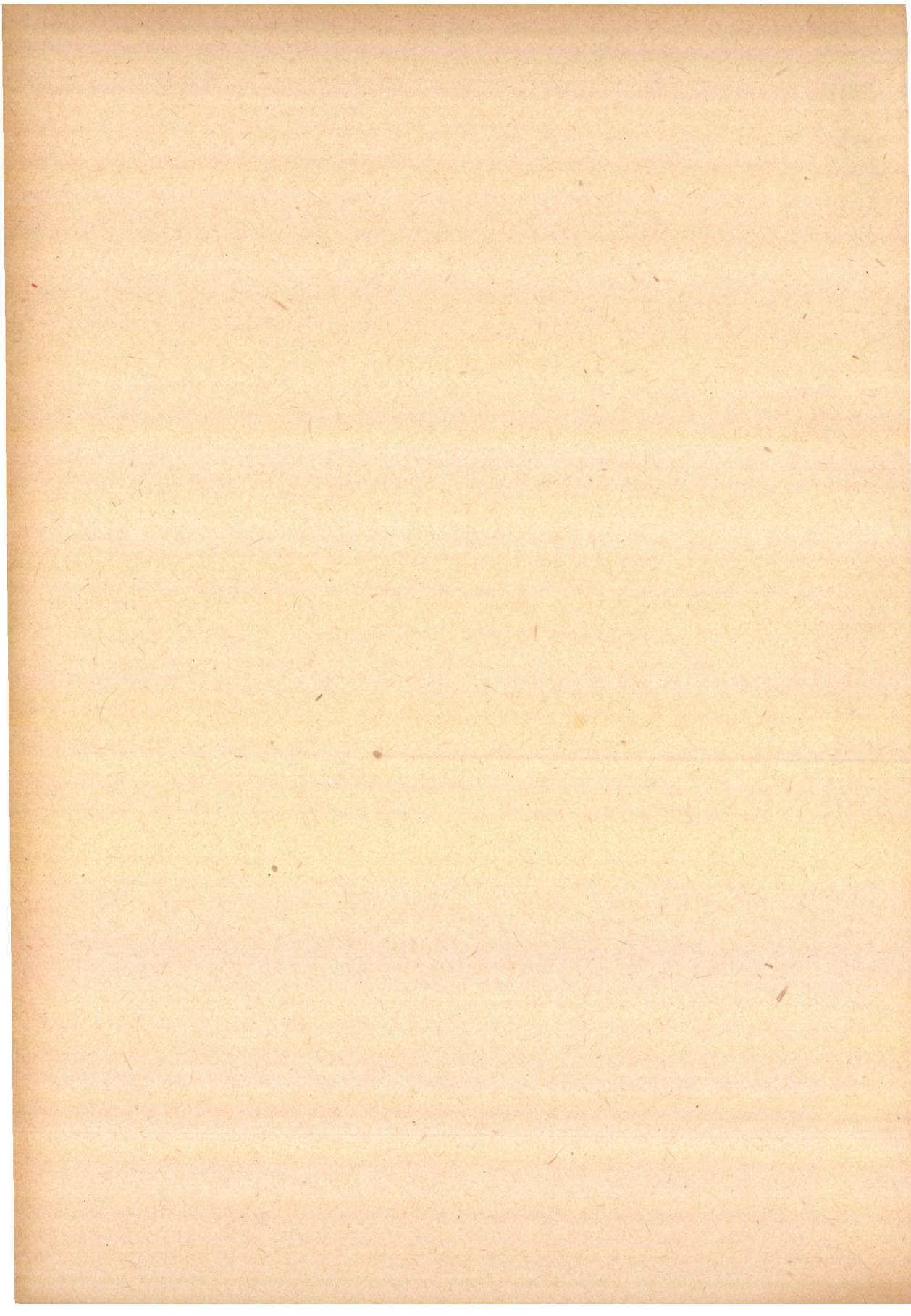
В настоящей статье автор доказывает, что число деревьев с n угловыми точками равно числу разложений данного цикла n -ой степени в произведение $n-1$ транспозиций (теорема 1).

Пусть C_n обозначает число разложений данного цикла n -ой степени в произведение $n-1$ транспозиций. Тогда $C_n = n^{n-2}$ (следствие).

Теорема 2 дает рекурсивное соотношение относительно C_n .

Теорема 4 определяет число разложений всех перестановок n -ой степени, содержащих k , цикл в произведение $n-k$ транспозиций.

Применение теоремы 4 в теории графов позволяет определить число графов с n угловыми точками, состоящих из k попарно различных деревьев без общих точек (теорема 5).



SOME REMARKS ON THE THEORY OF TREES

by

ALFRÉD RÉNYI

Introduction

Let T_n denote the number of trees of order n , i. e. of trees with n labelled vertices P_1, P_2, \dots, P_n . It has been proved by A. CAYLEY [1] that

$$(1) \quad T_n = n^{n-2} \quad (n = 1, 2, \dots).$$

Other proofs of (1) have been given by O. DZIOBEK [2] and G. PÓLYA [3]. The most elegant and simple proof is that given by PRÜFER [4]. The number T_n can be interpreted as the number of railway-nets connecting n given cities so that it should be possible to go from any city to any other along the net, and there are no superfluous connections. (The number of connections, i. e. the number of edges of a tree of order n is clearly $n - 1$). A. CAYLEY has treated also a more general problem. Let $T_{n,k}$ denote the number of graphs with $n + k - 1$ labelled vertices $P_1, P_2, \dots, P_{n+k-1}$ which consist of k disjoint trees as subgraphs and are such that the points P_1, P_2, \dots, P_k belong to different subgraphs. CAYLEY gives the following formula for $T_{n,k}$

$$(2) \quad T_{n,k} = k(n + k - 1)^{n-2}.$$

Clearly $T_{n,1} = T_n$, and thus (1) is a special case of (2).

The number $T_{n,k}$ can be interpreted as follows: If there are $n + k - 1$ cities, among which k cities are lying on the bank of the same river (or sea), then $T_{n,k}$ is the number of such minimal railway nets connecting some of the cities, that it is possible to go from any city to any other, by travelling by railway along the net and eventually travelling by ship from one of the cities lying at the river to an other such city. Thus the railway-nets in question will consist of k subnets, such that every city lying at the bank of the river belongs to an other subnet, and if the cities A and B belong to the same subnet, one can go by railway from A to B , but if A and B belong to different subnets, one can travel from A to B by going first by train to the harbour-city belonging to the same subnet as A , then by ship to the harbour-city belonging to the same net as B and finally by train to B .

CAYLEY gave (2) without proof and in the literature on (1) we did not find a proof of (2). In § 1 of the present paper we shall give a proof of (2), which is based on the same idea which has been used, e. g. by DZIOBEK in proving (1).¹ Our chief aim is however not this, but the solution of some other

¹ T. GALLAI kindly called my attention to the fact that (2) can be proved also by a modification of the method of PRÜFER.

related questions. In § 1 we shall determine the number $G_k(n)$ of all graphs with n labelled vertices consisting of k disjoint trees. We shall prove that

$$(3) \quad G_k(n) = \frac{1}{k!} \sum_{j=0}^k \left(-\frac{1}{2}\right)^j \binom{k}{j} \binom{n-1}{k+j-1} n^{n-k-j} \cdot (k+j)!.$$

Clearly we obtain from (3)

$$G_1(n) = T_n = n^{n-2}.$$

Thus (3) can also be considered as a generalization of (1). We obtain further by substituting $k=2$ and $k=3$ into (3)

$$(4) \quad G_2(n) = \frac{1}{2} n^{n-4} (n-1) (n+6)$$

and

$$(5) \quad G_3(n) = \frac{1}{8} n^{n-6} (n-1) (n-2) (n^2 + 13n + 60).$$

Similarly we can express $G_k(n)$ for other small values of k . It is easy to prove by means of (3) that for a fixed value of k and for $n \rightarrow +\infty$ we have

$$(6) \quad G_k(n) \sim \frac{n^{n-2}}{2^{k-1} \cdot (k-1)!}.$$

The formula (3) solves a problem left open in the paper of J. DÉNES [5], who proved that

$$(7) \quad G_k(n) = n! \sum_{\substack{\sum_{j=1}^n a_j = k \\ \sum_{j=1}^n j a_j = n}} \prod_{j=1}^n \frac{\binom{j-2}{j!}^{a_j}}{a_j!}$$

and asked whether the expression for $G_k(n)$ could be brought to a simpler form. (7) is a consequence of (1) and of the fact, that the partitioning of n points into a_1 sets having each one element, a_2 sets having each 2 elements, ..., a_j sets of j elements each ..., etc., where

$$\sum_{j=1}^n a_j = k$$

and

$$\sum_{j=1}^n j a_j = n$$

can be done in

$$n! \prod_{j=1}^n \frac{1}{a_j! (j!)^{a_j}}$$

essentially different ways. (Here two partitions are considered identical if there exists a one-to-one correspondence of the subsets of the two partitions so that corresponding subsets contain the same elements.)

In § 2 we consider an other problem, namely the determination of the number $T(n, r)$ of trees with n vertices which have exactly r end points. (A point belonging to a tree is called an end-point of the tree if only one edge of the tree contains the given point.) We shall prove that

$$(8) \quad T(n, r) = \frac{n!}{r!} \mathfrak{S}_{n-r}^{n-r}$$

where \mathfrak{S}_N^m denote the Stirling-numbers of the second kind defined by the identity

$$(9) \quad x^N = \sum_{m=1}^N \mathfrak{S}_N^m \cdot x(x-1) \dots (x-m+1).$$

By means of (8) we prove that if v_n denotes the number of end-points of a randomly chosen tree with n vertices, then the mean value of v_n is asymptotically n/e for $n \rightarrow +\infty$, further that the distribution of the random variable

$$\frac{v_n - \frac{n}{e}}{\frac{1}{e} \sqrt{(e-2)n}}$$

is tending for $n \rightarrow +\infty$ to the normal distribution with mean 0 and variance 1.

The author expresses his thanks to VERA T. SÓS and T. GALLAI for their valuable remarks.

§ 1. On the number of graphs consisting of trees

Our proof of (2) will be based on the identity

$$(1.1) \quad A_{n,k} = \sum_{\substack{k \\ \sum_{i=1}^k j_i = n}} \frac{n!}{j_1! j_2! \dots j_k!} j_1^{j_1-1} \cdot j_2^{j_2-1} \dots j_k^{j_k-1} = kn^{n-k} \prod_{h=1}^{k-1} (n-h)$$

where for $k=1$ the empty product means 1.

The identity (1.1) contains as a special case for $k=2$ the well-known identity

$$(1.2) \quad A_{n,2} = \sum_{j=1}^{n-1} \binom{n}{j} j^{j-1} (n-j)^{n-j-1} = 2n^{n-2}(n-1),$$

which has been proved e. g. in [2] and [5].

To prove (1.1) we shall need the identity

$$(1.3) \quad \left(\sum_{j=1}^{\infty} \frac{j^{j-1} x^j}{j!} \right)^k = \sum_{n=k}^{\infty} \frac{x^n \cdot kn^{n-k} \prod_{h=1}^{k-1} (n-h)}{n!}$$

where for $k = 1$ the empty product is to be replaced by 1. We shall prove (1.3) by induction on k . We shall start from the fact that

$$(1.4) \quad y = \sum_{j=1}^{\infty} \frac{j^{j-1} x^j}{j!}$$

is the Bürmann—Lagrange series for the inverse function $y = y(x)$ of the function

$$(1.5) \quad x = ye^{-y}$$

(see e. g. [6]). Now we obtain from (1.5)

$$(1.6) \quad ky^{k-1} y' = ky^{k-2} y' - \frac{ky^{k-1}}{x}$$

and integrating both sides of (1.6) from 0 to x we get

$$(1.7) \quad y^k = \frac{k}{k-1} y^{k-1} - k \int_0^x \frac{y^{k-1}(t)}{t} dt.$$

Now (1.3) is trivial for $k = 1$ and if it holds for $k-1$ instead of k , then by (1.7) it follows that it holds for k too. Thus (1.3) is proved for any k . (1.3) could be proved also directly by using a well-known generalized form of the Bürmann—Lagrange-series (see [6]).

Now evidently

$$(1.8) \quad \sum_{n=1}^{\infty} \frac{A_{n,k} x^n}{n!} = \left(\sum_{j=1}^{\infty} \frac{j^{j-1} x^j}{j!} \right)^k$$

and thus (1.3) and (1.8) imply (1.1).

Formula (1) follows easily by induction from (1.2), as was shown in [2]. For the sake of completeness we reproduce the proof.

Let us select for any $i = 1, 2, \dots, n-1$ an arbitrary subset having i elements, of the set of vertices P_1, \dots, P_n ; take an arbitrary tree connecting these points, and an arbitrary tree connecting the remaining $n-i$ points. If an arbitrary point of the first tree is connected with an arbitrary point of the second tree (which can be done in $i(n-i)$ different ways) we obtain a tree of order n . Evidently if these operations are effected in every possible way, every tree with vertices P_1, \dots, P_n is obtained $2(n-1)$ times. Thus we have

$$(1.9) \quad 2(n-1) T_n = \sum_{i=1}^{n-1} \binom{n}{i} i(n-i) T_i T_{n-i}.$$

Thus if $T_i = i^{i-2}$ holds for $i = 1, 2, \dots, n-1$, it holds by virtue of (1.2) for $i = n$ too. As clearly $T_1 = 1 = 1^{1-2}$, (1) follows by induction for all n .

As regards (2) it can be deduced from (1) by means of (1.1) as follows:

We have evidently

$$(1.10) \quad T_{n,k} = \sum_{\substack{k \\ \sum_{i=1}^k j_i = n-1}} \frac{(n-1)!}{j_1! j_2! \dots j_k!} T_{j_1+1} \cdot T_{j_2+1} \dots T_{j_k+1}$$

and thus

$$(1.11) \quad \prod_{j=1}^k (n+j-1) T_{n,k} = \sum_{\substack{k \\ \sum_{i=1}^k h_i = n+k-1}} \frac{(n+k-1)!}{h_1! h_2! \dots h_k!} h_1^{h_1-1} h_2^{h_2-1} \dots h_k^{h_k-1}$$

and therefore by (1.1)

$$(n+k-1)(n+k-2) \dots n \cdot T_{n,k} = k(n+k-1)^{n-1} \prod_{h=1}^{k-1} (n+k-1-h)$$

which, after dividing both sides by $(n+k-1)(n+k-2) \dots n$, gives evidently (2).

Now we pass to the proof of (3).

We start instead of (7), from the simpler formula

$$(1.12) \quad G_k(n) = \frac{1}{k!} \sum_{\substack{k \\ \sum_{i=1}^k j_i = n}} \frac{n!}{j_1! j_2! \dots j_k!} j_1^{j_1-2} j_2^{j_2-2} \dots j_k^{j_k-2}$$

(1.12) follows from (1) and the remark that all graphs formed with the vertices P_1, \dots, P_n and consisting of k trees can be obtained by forming all possible partitions into k subsets of the points P_1, \dots, P_n (two partitions being taken as identical if any two points belonging to the same subset in the first partition belong also to the same subset of the second partition) and by forming for each subset, independently of each other, all possible trees connecting the points of the subset with each other.

It follows from (1.12) [or from (7)] that

$$(1.13) \quad \sum_{n=1}^{\infty} \frac{G_k(n) x^n}{n!} = \frac{Y^k}{k!}$$

where Y is defined by the power series

$$(1.14) \quad Y = \sum_{j=1}^{\infty} \frac{j^{j-2} x^j}{j!}.$$

Now we have clearly

$$(1.15) \quad Y = \int_0^x \frac{y(t)}{t} dt$$

where y is the power series defined by (1.4); thus in view of (1.7) (for $k = 2$) we obtain

$$(1.16) \quad Y = y - \frac{y^2}{2}.$$

It follows that

$$(1.17) \quad \sum_{n=1}^{\infty} \frac{G_k(n) x^n}{n!} = \frac{1}{k!} \left(y - \frac{y^2}{2} \right)^k = \sum_{j=0}^k \frac{1}{k!} \binom{k}{j} \left(-\frac{1}{2} \right)^j y^{k+j}.$$

As by (1.8) we have

$$(1.18) \quad y^{k+j} = \sum_{n=1}^{\infty} \frac{A_{n,k+j} x^n}{n!}$$

it follows, taking (1.1) into account, that

$$(1.19) \quad G_k(n) = \frac{1}{k!} \sum_{j=0}^k \binom{k}{j} \left(-\frac{1}{2} \right)^j (k+j) n^{n-k-j} \prod_{h=1}^{k+j-1} (n-h).$$

Thus (3) is proved.

Let us calculate the coefficient of the two highest power of n on the right of (1.19). Putting

$$G_k(n) = \sum_{h=1}^{k+j} G_{n,k}^{(h)} n^{n-h}$$

we obtain by some calculation

$$G_{n,k}^{(1)} = 0 \quad \text{and} \quad G_{n,k}^{(2)} = \frac{1}{2^{k-1} \cdot (k-1)!}.$$

Thus we have for any fixed k

$$(1.20) \quad \lim_{n \rightarrow +\infty} \frac{G_k(n)}{n^{n-2}} = \frac{1}{2^{k-1} \cdot (k-1)!}.$$

Of course $G_1(n) = T_n = n^{n-2}$. Other trivial special cases of (1.19) are

$$(1.21) \quad \begin{aligned} G_n(n) &= 1, \\ G_{n-1}(n) &= \binom{n}{2}, \\ G_{n-2}(n) &= 3 \binom{n+1}{4}. \end{aligned}$$

Let us put

$$(1.22) \quad H(n) = \sum_{k=1}^n G_k(n) .$$

Clearly $H(n)$ is the number of graphs of order n which are the unions of trees, i. e. of all graphs which do not contain circles (closed paths). From (1.17) we obtain easily, putting $H(0) = 1$,

$$(1.23) \quad \sum_{n=0}^{\infty} \frac{H(n) x^n}{n!} = e^{y - \frac{y^2}{2}} ,$$

As

$$e^{-xy - \frac{y^2}{2}} = \sum_{r=0}^{\infty} H_r(x) y^r$$

where

$$H_r(x) = \frac{1}{r!} e^{\frac{x^2}{2}} \frac{d^r}{dx^r} \left(e^{-\frac{x^2}{2}} \right)$$

is the r -th Hermite polynomial, we obtain

$$(1.24) \quad \sum_{n=0}^{\infty} \frac{H(n) x^n}{n!} = \sum_{r=0}^{\infty} H_r(-1) y^r$$

and thus, taking again (1.1) into account, we obtain

$$(1.25) \quad H(n) = \sum_{r=1}^n H_r(-1) r n^{n-r} \prod_{h=1}^{r-1} (n-h) .$$

As regards the asymptotic behaviour of $H(n)$ it follows from (1.20) easily that

$$(1.26) \quad \lim_{n \rightarrow +\infty} \frac{H(n)}{n^{n-2}} = \sqrt{e} .$$

§ 2. On the number of end-points of a random tree

Let us consider a tree whose vertices are the points P_1, \dots, P_n . The point P_i is called an end-point of the tree, if there is only one edge of the tree connecting P_i with some other point P_j . The number of end-points of a tree of order n may have the value 2, 3, ..., $n-1$. Let $T(n, k)$ denote the number of trees with the (labelled) vertices P_1, \dots, P_n which have exactly k end-points ($k = 2, 3, \dots, n-1$). We shall prove first the formula (8).

Our proof utilizes the method by which PRÜFER [4] proved formula (1). PRÜFER's method consists in that he establishes a one-to-one correspondence between all trees with vertices P_1, \dots, P_n and all $(n-2)$ -tuples of integers

$(s_1, s_2, \dots, s_{n-2})$ where each s_j can take the values $1, 2, \dots, n$. This correspondence is obtained as follows: let us remove from the tree the endpoint P_{i_1} with the least index, and let s_1 denote the index of the (unique) point which is connected by an edge with P_{i_1} . Repeating the same operation with the remaining tree of order $n-1$ we obtain s_2, s_3 , etc. until only a single edge remains. PRÜFER has shown that the sequence $(s_1, s_2, \dots, s_{n-2})$ determines the tree uniquely. Now clearly to the trees with exactly k endpoints there correspond sequences $(s_1, s_2, \dots, s_{n-2})$ in which exactly $n-k$ of the numbers $1, 2, \dots, n$ occur at least once. As the failing numbers can be chosen in $\binom{n}{k}$ different ways, and the number of $(n-2)$ -tuples formed from $n-k$ symbols which contain each of these $n-k$ symbols at least once is equal to $(n-k)! \mathfrak{S}_{n-k}^{n-k}$ (see e. g. [7]), it follows that

$$(2.1) \quad T(n, k) = \frac{n!}{k!} \mathfrak{S}_{n-k}^{n-k}.$$

According to the well-known recursion formula $\mathfrak{S}_n^k = k \mathfrak{S}_{n-1}^k + \mathfrak{S}_{n-1}^{k-1}$ (see e. g. [7], p. 169) it follows from (2.1) that

$$(2.2) \quad \frac{k}{n} T(n, k) = k T(n-1, k) + (n-k) T(n-1, k-1).$$

Conversely (2.1) can be deduced from (2.2). The recursion formula (2.2) can be proved by a direct combinatorial argument, as has been remarked by VERA T. SÓS.

Let us calculate now the mean value and variance of the number of endpoints of a random tree of order n ; here and in what follows if we speak about a random tree of order n we mean by this that we select at random a tree of order n with given vertices P_1, \dots, P_n so that all n^{n-2} possible trees are equiprobable. With this definition the number v_n of endpoints of a random tree is a random variable. Let us denote by $\mathbf{M}\{v_n\}$ resp. $\mathbf{D}^2\{v_n\}$ the mean value resp. the variance of v_n . Now according to the definition of Stirling's numbers of the second kind [see (9)] from (2.1) we have

$$(2.3) \quad \sum_{k=2}^{n-1} T(n, k) \frac{\binom{x}{n-k}}{\binom{n}{n-k}} = x^{n-2}.$$

The formula (2.3) can be considered as the (factorial) generating function of the sequence $T(n, k)$. Substituting now $x = n-1$ into (2.3) we obtain

$$\sum_{k=2}^{n-1} k T(n, k) = n(n-1)^{n-2}$$

and thus

$$(2.4) \quad \mathbf{M}\{v_n\} = n \left(1 - \frac{1}{n}\right)^{n-2}.$$

Therefore we have

$$(2.5) \quad \lim_{n \rightarrow +\infty} \frac{\mathbf{M}\{v_n\}}{n} = \frac{1}{e}.$$

This can be expressed by saying that a random tree of order n has in the average approximately n/e endpoints.

To calculate the variance of v_n let us substitute $x = n - 2$ into (2.2). We obtain

$$(2.6) \quad \sum_{k=2}^{n-1} k(k-1) T(n, k) = n(n-1)(n-2)^{n-2}$$

and thus

$$(2.7) \quad \sum_{k=2}^{n-1} k^2 T(n, k) = n(n-1)(n-2)^{n-2} + n(n-1)^{n-2}.$$

It follows that

$$(2.8) \quad \mathbf{D}^2\{v_n\} = n(n-1) \left(1 - \frac{2}{n}\right)^{n-2} + n \left(1 - \frac{1}{n}\right)^{n-2} - n^2 \left(1 - \frac{1}{n}\right)^{2n-4}.$$

This implies

$$(2.9) \quad \lim_{n \rightarrow +\infty} \frac{\mathbf{D}^2\{v_n\}}{n} = \frac{e-2}{e^2}.$$

Thus the variance of v_n is asymptotically equal to $n(e-2)/e^2$.

Now we shall prove that v_n is in the limit normally distributed; more exactly we shall prove

$$(2.10) \quad \lim_{n \rightarrow +\infty} \mathbf{P} \left\{ \frac{v_n - \frac{n}{e}}{\frac{1}{e} \sqrt{n(e-2)}} < x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$$

where $\mathbf{P}\{\dots\}$ denotes the probability of the event in the brackets.

To prove (2.10), by a well-known theorem of probability theory, it suffices to show that the characteristic function of

$$\frac{v_n - \frac{n}{e}}{\frac{1}{e} \sqrt{n(e-2)}}$$

tends to the characteristic function of the normal distribution, that is for every real t we have

$$(2.11) \quad \lim_{n \rightarrow +\infty} \sum_{k=2}^{n-1} \frac{T(n, k)}{n^{n-2}} e^{\frac{it(ek-n)}{n(e-2)}} = e^{-\frac{t^2}{2}}.$$

Now (2.11) can be obtained as follows: Let us substitute into (2.2) $x = n - it\sqrt{n}$. We obtain

$$(2.12) \quad \sum_{k=2}^{n-1} \frac{T(n, k)}{n^{n-2}} \prod_{j=k+1}^n \left(1 - \frac{it\sqrt{n}}{j}\right) = \left(1 - \frac{it}{\sqrt{n}}\right)^{n-2}.$$

If

$$\left|k - \frac{n}{e}\right| < n^\beta$$

where $1/2 < \beta < 1$, we have

$$(2.13) \quad \log \prod_{j=0}^{n-k-1} \left(1 - \frac{it\sqrt{n}}{n-j}\right) = -it\sqrt{n} + \frac{ite}{\sqrt{n}} \left(k - \frac{n}{e}\right) + \frac{t^2(e-1)}{2} + o(1).$$

By the inequality of Chebysheff

$$(2.14) \quad \sum_{\left|k - \frac{n}{e}\right| > n^\beta} \frac{T(n, k)}{n^{n-2}} = O\left(\frac{1}{n^{2\beta-1}}\right)$$

and according to Stirling's formula for complex argument the factor

$$\prod_{j=k+1}^n \left(1 - \frac{it\sqrt{n}}{j}\right)$$

remains bounded; thus it follows that

$$(2.15) \quad \lim_{n \rightarrow +\infty} \left| \left(1 - \frac{it}{\sqrt{n}}\right)^{n-2} e^{it\sqrt{n}} - e^{\frac{t^2(e-1)}{2}} \sum_{k=2}^{n-1} \frac{T(n, k)}{n^{n-2}} e^{\frac{ite}{\sqrt{n}} \left(k - \frac{n}{e}\right)} \right| = 0.$$

Now we have

$$(2.16) \quad \lim_{n \rightarrow +\infty} \left(1 - \frac{it}{\sqrt{n}}\right)^{n-2} e^{it\sqrt{n}} = e^{\frac{t^2}{2}}.$$

Thus it follows that

$$(2.17) \quad \lim_{n \rightarrow +\infty} \sum_{k=2}^n \frac{T(n, k)}{n^{n-2}} e^{\frac{ite}{\sqrt{n}} \left(k - \frac{n}{e}\right)} = e^{-\frac{t^2(-2)}{2}}.$$

Substituting $\frac{t}{\sqrt{e-2}}$ instead of t , (2.11) and therefore (2.10) follows.

(Received 30 October, 1958.)

Added in proof 5 May, 1959.

In a recent paper [8] L. E. CLARKE gave a new proof of CAYLEY's formula (1). CLARKE proves first that if $C_{n,k}$ denotes the number of trees with n labelled vertices P_1, P_2, \dots, P_n such that exactly k edges end at the vertex P_n , then we have

$$(2.18) \quad C_{n,k} = \binom{n-2}{k-1} (n-1)^{n-k-1} \quad (k = 1, 2, \dots, n-1).$$

CLARKE proves (2.18) by giving a recursion formula for $C_{n,k}$, and deduces (1) from (2.18) by remarking that

$$(2.19) \quad T_n = \sum_{k=1}^{n-1} C_{n,k}.$$

It should be mentioned that (2.18) is a simple consequence of CAYLEY's formula (2). As a matter of fact, the k vertices which are connected with P_n can be chosen in $\binom{n-1}{k}$ ways among the vertices P_1, P_2, \dots, P_{n-1} and after this choice has been made, there remain still $T_{n-k,k}$ possible choices to form the tree in question. Thus we have

$$(2.20) \quad C_{n,k} = \binom{n-1}{k} T_{n-k,k}$$

and therefore by (2)

$$(2.21) \quad C_{n,k} = \binom{n-1}{k} k(n-1)^{n-k-2} = \binom{n-2}{k-1} (n-1)^{n-k-1}.$$

in accordance with (2.18).

REFERENCES

- [1] CAYLEY, A.: „A theorem on trees”. *Quarterly Journal of Pure and Applied Mathematics* **23** (1889) 376—378. See also *The Collected Papers of A. Cayley*, Cambridge, 1897, Vol. 13., pp. 26—28.
- [2] DZIOBEK, O.: „Eine Formel der Substitutionstheorie”, *Sitzungsberichte der Berliner Mathematischen Gesellschaft* **17** (1917) 64—67.
- [3] PÓLYA, G.: „Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen.” *Acta Mathematica (Uppsala)* **68** (1937) 145—255.
- [4] PRÜFER, A.: „Neuer Beweis eines Satzes über Permutationen” *Archiv für Mathematik und Physik* **27** (1918) 142—144.
- [5] DÉNES, J.: „The representation of a permutation as the product of a minimal number of transpositions and its connection with the theory of graphs”. *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* **4** (1959) 63—71.
- [6] PÓLYA, G.—SZEGŐ, G.: *Aufgaben und Lehrsätze aus der Analysis*. Springer, Berlin, 1925, Vol. I., p. 125.
- [7] JORDAN, CH.: *Calculus of finite differences*. Budapest, 1939, p. 177.
- [8] CLARKE, L. E.: „On Cayley's formula for counting trees”, *Journal of the London Mathematical Society* **33** (1958) 411—475.

MEGJEGYZÉSEK A „FÁK” ELMÉLETÉHEZ

RÉNYI A.

Kivonat

„Fa” alatt CAYLEY [1] nyomán körnélküli összefüggő gráfot értünk. CAYLEY bebizonyította, hogy az n (számozott) szögponttal bíró különböző fák száma

$$(1) \quad T_n = n^{n-2},$$

míg azon $n + k - 1$ (számozott) súlyponttal bíró gráfok száma, amelyek k idegen fából állnak oly módon, hogy k kijelölt pont különböző fákhoz tartozik,

$$(2) \quad T_{n,k} = k(n + k - 1)^{n-2}.$$

Az (1) képletre CAYLEYN kívül O. DZIOBEK [2], PÓLYA GYÖRGY [3] és mások adtak más bizonyításokat. (1) legegyszerűbb és legelegánsabb bizonyítása PRÜFERTől származik [4]. Az 1. §-ban rámutat a szerző, hogy a (2) összefüggés (amely CAYLEYNél bizonyítás nélkül szerepel) egyszerűen bebizonyítható DZIOBEK módszerével.

Jelölje $G_k(n)$ azon n (számozott) szögpontú különböző gráfok számát, amelyek k idegen fából állnak, (elejtve a CAYLEY által tett megszorítást, hogy k adott pont különböző fákhoz tartozzék). Az 1 §-ban bebizonyítja a szerző, hogy

$$(3) \quad G_k(n) = \frac{1}{k!} \sum_{j=0}^k \binom{k}{j} (k+j)! \left(\frac{n-1}{k+j-1} \right) \left(-\frac{1}{2} \right)^j n^{n-k-j}.$$

A (3) képlet segítségével $G_k(n)$ értékére, ha k rögzítve van és $n \rightarrow +\infty$ aszimptotikus képlet adható meg.

A 2. §. azzal a kérdéssel foglalkozik, hogy hány olyan n szögpontú fa van, amelynek pontosan r végpontja van. Ezt a számot $T(n, r)$ -rel jelölve PRÜFER módszerével kimutatható, hogy

$$(8) \quad T(n, r) = \frac{n!}{r!} \mathfrak{S}_{n-2}^{n-r},$$

ahol \mathfrak{S}_N^m másodfajú Stirling-számokat jelöli, vagyis a \mathfrak{S}_N^m számok az

$$(9) \quad x^N = \sum_{m=1}^N \mathfrak{S}_N^m x(x-1) \dots (x-m+1)$$

összefüggés által vannak definiálva.

A (8) explicit képlet segítségével a szerző kimutatja, hogy egy taláalomra választott n (számozott) szögpontú fa végpontjainak száma határértékben normális eloszlású n/e várható értékkel és $\sqrt{(e-2)n/e}$ szórással, ha $n \rightarrow +\infty$.

ЗАМЕЧАНИЯ К ТЕОРИИ «ДЕРЕВЬЕВ»

A. RÉNYI

Резюме

Следуя CAYLEY [1] связанный граф, не содержащий окружностей называется «деревом». CAYLEY доказал, что число деревьев, имеющих n (нумерированных) угловых точек, равно

$$(1) \quad T_n = n^{n-2}$$

в то время, как число графов с $n+k-1$ (нумерированными) точками состоящих из k деревьев так, что k указанных точек принадлежат различным деревьям, равно

$$(2) \quad T_{n,k} = k(n+k-1)^{n-2}.$$

Для формулы (1) кроме CAYLEY другие доказательства дали также O. DZIOBEK [2] и G. PÓLYA [3]. Самое простое и элегантное доказательство (1)

принадлежит PRÜFER-у [4]. В 1. §-е автор показывает, что соотношение (2) (которое у CAULEY приведено без доказательства) также может быть доказано, методом DZIOBEK-а.

Обозначим через $G_k(n)$ число тех различных графов с n (нумерированными) угловыми точками, которые состоят из k деревьев (отказавшись от условия CAULEY, согласно которому k данных точек принадлежат различным деревьям). В 1. §-е автор доказывает, что

$$(3) \quad G_k(n) = \frac{1}{k!} \sum_{j=0}^k \binom{k}{j} (k+j)! \left(\begin{matrix} n-1 \\ k+j-1 \end{matrix} \right) \left(-\frac{1}{2} \right)^j n^{n-k-j}.$$

С помощью формулы (3) для значения $G_k(n)$ при фиксированном k и $n \rightarrow \infty$ может быть дана асимптотическая формула.

§ 2. занимается следующим вопросом: сколько существует таких деревьев с n угловыми точками, которые имеют точно r конечных точек. Обозначая это число через $T(n, r)$, методом PRÜFER-а можно доказать, что

$$(8) \quad T_{n,r} = \frac{n!}{r!} \mathfrak{S}_{n-2}^{n-r}$$

где \mathfrak{S}_N^m обозначают числа Stirling-а второго рода, т. е. определяются соотношением

$$(9) \quad x^N = \sum_{m=1}^N \mathfrak{S}_N^m x(x-1) \dots (x-m+1).$$

С помощью явной формулы (8) автор доказывает, что число конечных точек случайно выбранного дерева с n (нумерированными) угловыми точками в пределе имеет нормальное распределение с математическим ожиданием n/e и дисперсией $\frac{(e-2)n}{e^2}$ если $n \rightarrow \infty$.

A NOTE ON A PROBLEM OF RÉNYI

D. C. HANDSCOMB

(Christ Church, Oxford, England)

A. RÉNYI¹ recently propounded the following problem. An $n \times n$ matrix $\mathbf{X} = \{x_{ij}\}$, whose elements are all real and nonnegative, has to satisfy the conditions $\{\sum_j x_{ij} = g_i\}$ and $\{\sum_i x_{ij} = h_j\}$, and also has to minimize the linear form $\sum_{ij} c_{ij} x_{ij}$. Under these conditions, a certain number of the elements of \mathbf{X} will be zero. If more than one matrix satisfies the conditions, we select one with as many zero elements as possible. If now we consider every possible set of values for $\{g_i\}$, $\{h_j\}$ and $\{c_{ij}\}$, the resulting matrices \mathbf{X} will have various numbers of zero elements; let $f(n)$ be the smallest such number for a given value of n . It is the form of $f(n)$ which is required.

It is easily shown that $f(n) \leq (n-1)^2$. Consider for instance the situation where the linear form to be minimized consists of the sum of all the elements of \mathbf{X} except for the first row and the first column, and where $g_1 = h_1 = n$, and $g_i = h_j = 1$ for i and $j \neq 1$. In fact, as will be shown below, $f(n) = (n-1)^2$.

Treat the elements of \mathbf{X} as coordinates of a point in euclidean space of n^2 dimensions. The fixing of g_i and h_j impose $2n-1$ independent conditions, which restrict the point to a linear subspace of $(n-1)^2$ dimensions. Now the conditions $\{x_{ij} \geq 0\}$ have the added effect of restricting the point to a region \mathcal{R} of this subspace, where \mathcal{R} is convex and is bounded by surfaces of the form $x_{ij} = 0$. This being so, it is clear that the minimum over \mathcal{R} of the form $\sum c_{ij} x_{ij}$ is attained either at a vertex of \mathcal{R} or over a set which contains vertices of \mathcal{R} . But every vertex of \mathcal{R} must lie on at least $(n-1)^2$ of the bounding surfaces of \mathcal{R} , and thus at least this number of the elements of \mathbf{X} are zero at a vertex. Therefore $f(n) \geq (n-1)^2$. Combining this with the above, we have $f(n) = (n-1)^2$.

Similarly it may be proved that the corresponding number for $n \times m$ matrices is $(n-1)(m-1)$, and similar results hold for arrays of more than 2 dimensions.

From the above it is possible to deduce the result, proved previously by other means by BIRKHOFF [1] and by HAMMERSLEY and MAULDON [2], that the set of $n \times n$ doubly-stochastic matrices is the convex hull of the set

¹ In a lecture at the Colloquium on Monte-Carlo methods, held in Balatonvilágos, 23rd September, 1958. A summary of this lecture is given in "Matematikai Lapok" 9 (1958) p. 353-354.

of $n \times n$ permutation matrices. \mathbf{X} is doubly stochastic if $g_i = h_j = 1$. But if \mathbf{X} is a vertex of \mathcal{Q} it has at most $n^2 - (n-1)^2 = 2n-1$ non-zero elements. Therefore at least one row of \mathbf{X} has only 1 non-zero element, which must have the value 1, and which must therefore also be the only non-zero element in its column. Delete this row and column and proceed by induction.

This note is the outcome of a private communication from Dr. A. RÉNYI, and of discussion with Mr. J. M. HAMMERSLEY, to both of whom I am grateful, as I am also to the University of Illinois who are at present supporting me.

(Received. January, 31 1959)

REFERENCES

- [1] BIRKHOFF, G.: *Rev. Univ. Tucuman A*, 5 (1946), 147—151.
 [2] HAMMERSLEY J. M. and MAULDON J. G.: *Proc Camb. Phil. Soc.* 52 (1956) 461—481

RÉNYI ALFRÉD EGY PROBLÉMÁJÁRÓL

D. C. HANDSCOMB

Kivonat

A dolgozat egy RÉNYI ALFRÉD által felvetett problémát old meg, amelyben bebizonyítja a következő tételt. Vizsgáljuk azokat az $\mathbf{X} = \{x_{ij}\}$ $n \times n$ -es nemnegatív elemű mátrixokat, amelyek elemei eleget tesznek a

$$\sum_j x_{ij} = g_i \quad (i = 1, 2, \dots, n), \quad \sum_i x_{ij} = h_j \quad (j = 1, 2, \dots, n)$$

feltételeknek; e mátrixok közül tekintsük azokat, amelyek a $\sum_i \sum_j c_{ij} x_{ij}$ lineáris alak értékét minimalizálják. Itt g_i és h_j tetszőleges előírt pozitív számok, a c_{ij} számok pedig tetszőleges valós számok. Az ilyen mátrixok között van olyan, amelyben a zérusok száma $\geq (n-1)^2$, és ez az eredmény nem javítható.

ОБ ОДНОЙ ПРОБЛЕМЕ А. RÉNYI

D. C. HANDSCOMB

Резюме

Работа решает одну проблему А. RÉNYI, доказывая следующую теорему. Рассмотрим матрицы $\mathbf{X} = \{x_{ij}\}$ с $n \times n$ неотрицательными элементами, удовлетворяющими условиям

$$\sum_j x_{ij} = g_i \quad (i = 1, 2, \dots, n), \quad \sum_i x_{ij} = h_j \quad (j = 1, 2, \dots, n);$$

среди этих матриц будем изучать те, которые минимизируют значение линейной формы $\sum_i \sum_j c_{ij} x_{ij}$. Здесь g_i и h_j любые фиксированные положительные числа, а c_{ij} любые вещественные числа. Среди этих матриц есть такая, в которой число нулей $\geq (n-1)^2$, и этот результат не может быть улучшен.

COMPLETELY CONTINUOUS OPERATORS WITH UNIFORMLY BOUNDED ITERATES

by

BÉLA SZ. NAGY

I have proved in 1945 the following theorem [1]. Let \mathbf{T} be an invertible operator¹ in (real or complex) euclidean space of a finite or infinite number of dimensions, such that all the iterates of \mathbf{T} and of \mathbf{T}^{-1} have a common bound, i. e.

$$\|\mathbf{T}^n\| \leq M \quad (n = 0, \pm 1, \pm 2, \dots).$$

Then \mathbf{T} is similar to a unitary operator. More precisely, there exists a self-adjoint invertible operator \mathbf{Q} such that

$$\mathbf{Q}\mathbf{T}\mathbf{Q}^{-1}$$

is *unitary*.

If \mathbf{T} is not invertible, but all its direct iterates $\mathbf{T}, \mathbf{T}^2, \mathbf{T}^3, \dots$ have a common bound, then, obviously, \mathbf{T} cannot be similar to any unitary operator. All that one may surmise in this case is that \mathbf{T} is similar to a *contraction*, i. e. to an operator \mathbf{T}' with $\|\mathbf{T}'\| \leq 1$.

In this note we shall show that this is in fact true at least if the space is complex and \mathbf{T} is completely continuous.² Since in finite dimensional euclidean space every operator is completely continuous, this settles our problem for these spaces in the positive sense. For non completely continuous operators in infinite dimensional euclidean space, however, the problem remains as yet open.

Thus, we shall prove the following

Theorem. *If \mathbf{T} is a completely continuous operator on a complex euclidean space \mathcal{R} of a finite or infinite number of dimensions, such that*

$$(1) \quad \|\mathbf{T}^n\| \leq M \quad (n = 0, 1, 2, \dots),$$

then \mathbf{T} is similar to a contraction; more precisely, there exists an invertible self-

¹ By an "operator" we shall mean always a bounded linear operator, which is everywhere defined in the space; an operator will be said "invertible" if it has an everywhere defined, bounded inverse. For any operator \mathbf{T} the norm $\|\mathbf{T}\|$ is defined as usual by the least upper bound of the length of the vector $\mathbf{T}\mathbf{x}$ when the vector \mathbf{x} varies on the unit sphere.

² I. e. if \mathbf{T} transforms any bounded set of vectors into a compact set; see [2] No. 85.

adjoint operator \mathbf{Q} on \mathcal{R} such that the operator $\mathbf{T}' = \mathbf{Q}\mathbf{T}\mathbf{Q}^{-1}$ is a contraction, i. e.

$$\|\mathbf{T}'\| \leq 1.$$

Proof. We make use of the „spectral radius theorem” (due to I. GELFAND in its general form for normed rings). For operators (on an arbitrary Banach space) it states that

$$\lim_{k \rightarrow \infty} \|\mathbf{T}^k\|^{\frac{1}{k}} = r(\mathbf{T})$$

where $r(\mathbf{T})$ is the „spectral radius” of the operator \mathbf{T} , i. e. the radius of the smallest closed disk in the plane of complex numbers, with centre at the point 0, which contains all the points of the spectrum of \mathbf{T} (see [2] No. 149).

Thus the condition (1) implies that $r(\mathbf{T}) \leq 1$. Since \mathbf{T} is completely continuous, its spectrum consists of a denumerable set of complex numbers having no accumulation point different from 0. Let $\lambda_1, \dots, \lambda_n$ be those points of the spectrum of \mathbf{T} (if any) which lie on the unit circle. The other points of the spectrum form a set σ_0 lying in some concentric circle of radius *smaller* than 1. To the decomposition

$$\sigma = \sigma_0 \cup \{\lambda_1\} \cup \dots \cup \{\lambda_n\}$$

of the spectrum σ of \mathbf{T} into isolated parts, there corresponds a decomposition

$$\mathbf{I} = \mathbf{P}_0 + \mathbf{P}_1 + \dots + \mathbf{P}_n,$$

of the identity operator \mathbf{I} into a sum of idempotent operators with

$$\mathbf{P}_i \mathbf{P}_k = \mathbf{O} \quad (i \neq k);$$

this decomposition is characterized by the property that \mathbf{T} is permutable with all \mathbf{P}_i ’s, and when considered as an operator on the subspace

$$\mathcal{R}_i = \mathbf{P}_i \mathcal{R},$$

\mathbf{T} has the spectrum σ_0 if $i = 0$ and the one-point spectrum $\{\lambda_i\}$ if $i \neq 0$ (see [2] No. 148); moreover, as another consequence of complete continuity, the subspaces \mathcal{R}_i ($i \neq 0$) are all finite dimensional.

Consider \mathbf{T} as an operator on the finite dimensional space \mathcal{R}_i ($i \neq 0$); we shall show that its matrix (in an arbitrary base of \mathcal{R}_i) has no non-linear elementary divisors. In the contrary case there should exist non-zero vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r$ in \mathcal{R}_i , forming a „Jordan-chain” of length $r \geq 2$, i. e. such that

$$\mathbf{T}\mathbf{x}_1 = \lambda_i \mathbf{x}_1, \quad \mathbf{T}\mathbf{x}_2 = \lambda_i \mathbf{x}_2 + \mathbf{x}_1, \quad \mathbf{T}\mathbf{x}_3 = \lambda_i \mathbf{x}_3 + \mathbf{x}_2, \dots, \quad \mathbf{T}\mathbf{x}_r = \lambda_i \mathbf{x}_r + \mathbf{x}_{r-1}.$$

From the first two of these relations it results by iteration

$$\mathbf{T}^k \mathbf{x}_2 = \lambda_i^k \mathbf{x}_2 + k \lambda_i^{k-1} \mathbf{x}_1 \quad (k = 1, 2, \dots).$$

Now by (1)

$$\|\mathbf{T}^k \mathbf{x}_2\| \leq M \|\mathbf{x}_2\|;$$

since $|\lambda_i| = 1$, this implies

$$\|\lambda_i \mathbf{x}_2 + k \mathbf{x}_1\| \leq M \|\mathbf{x}_2\|.$$

This is impossible, for $\|\lambda_i \mathbf{x}_2 + k \mathbf{x}_1\| \rightarrow \infty$ if $k \rightarrow \infty$. Thus there can be no non-linear elementary divisors, and since the only point of the spectrum of

\mathbf{T} on \mathcal{R}_i is λ_i , we conclude that for all $\mathbf{x} \in \mathcal{R}_i$, $\mathbf{T}\mathbf{x} = \lambda_i \mathbf{x}$. We have shown, in other words, that

$$(2) \quad \mathbf{T}\mathbf{P}_i = \mathbf{P}_i\mathbf{T} = \lambda_i \mathbf{P}_i \quad (i = 1, 2, \dots, n).$$

Consider now the operator $\mathbf{T}_0 = \mathbf{T}\mathbf{P}_0 = \mathbf{P}_0\mathbf{T}$. On \mathcal{R}_0 it coincides with \mathbf{T} , thus it has there the spectrum σ_0 , while on the complementary subspace $(\mathbf{I} - \mathbf{P}_0)\mathcal{R}$ it is identically 0. The spectrum of \mathbf{T}_0 is thus equal to σ_0 or $\sigma_0 \cup \{0\}$; in any case, its spectral radius $r(\mathbf{T}_0)$ is < 1 . Let p be a number between $r(\mathbf{T}_0)$ and 1. Then we have, by the spectral radius theorem,

$$\lim_{k \rightarrow \infty} \|\mathbf{T}_0^k\|^{\frac{1}{k}} < p < 1.$$

Thus, for sufficiently large values of k ,

$$\|\mathbf{T}_0^k\| < p^k,$$

and consequently

$$\|(\mathbf{T}_0^k)^* \mathbf{T}_0^k\| < p^{2k}.$$

This implies that the operator series

$$\sum_{k=1}^{\infty} (\mathbf{T}_0^k)^* \mathbf{T}_0^k$$

converges in the operator norm.

Consider the operator

$$\mathbf{A} = \sum_{i=0}^n \mathbf{P}_i^* \mathbf{P}_i + \sum_{k=1}^{\infty} (\mathbf{T}_0^*)^k \mathbf{T}_0^k;$$

it is selfadjoint, and we have for any $\mathbf{x} \in \mathcal{R}$

$$\begin{aligned} (\mathbf{A}\mathbf{x}, \mathbf{x}) &= \sum_{i=0}^n \|\mathbf{P}_i \mathbf{x}\|^2 + \sum_{k=1}^{\infty} \|\mathbf{T}_0^k \mathbf{x}\|^2 \geq \sum_{i=0}^n \|\mathbf{P}_i \mathbf{x}\|^2 \geq \frac{1}{n+1} \left(\sum_{i=0}^n \|\mathbf{P}_i \mathbf{x}\| \right)^2 > \\ &\geq \frac{1}{n+1} \left\| \sum_{i=0}^n \mathbf{P}_i \mathbf{x} \right\|^2 = \frac{1}{n+1} \|\mathbf{x}\|^2, \end{aligned}$$

thus

$$(3) \quad \mathbf{A} \geq \frac{1}{n+1} \mathbf{I}.$$

Let \mathbf{Q} be the positive selfadjoint square-root of \mathbf{A} ; from (3) it follows that \mathbf{Q} has the positive lower bound $(n+1)^{-1/2}$. This assures that \mathbf{Q} is invertible and $\|\mathbf{Q}^{-1}\| \leq (n+1)^{1/2}$.

Now we shall calculate $\mathbf{T}^* \mathbf{A} \mathbf{T}$. For $k \geq 1$ we have

$$\mathbf{T}_0^k \mathbf{T} = \mathbf{T}_0^{k-1} \mathbf{T} \mathbf{P}_0 \mathbf{T} = \mathbf{T}_0^{k-1} \mathbf{T} \mathbf{P}_0^2 \mathbf{T} = \mathbf{T}_0^{k-1} \mathbf{T} \mathbf{P}_0 \mathbf{T} \mathbf{P}_0 = \mathbf{T}_0^{k+1}$$

and

$$\mathbf{T}^* (\mathbf{T}_0^k)^* \mathbf{T}_0^k \mathbf{T} = (\mathbf{T}_0^k \mathbf{T})^* (\mathbf{T}_0^k \mathbf{T}) = (\mathbf{T}_0^{k+1})^* \mathbf{T}_0^{k+1},$$

thus

$$\mathbf{T}^* \left(\sum_{k=1}^{\infty} (\mathbf{T}_0^k)^* \mathbf{T}_0^k \right) \mathbf{T} = \sum_{k=2}^{\infty} (\mathbf{T}_0^k)^* \mathbf{T}_0^k .$$

Further we have

$$\mathbf{T}^* \mathbf{P}_0^* \mathbf{P}_0 \mathbf{T} = \mathbf{T}_0^* \mathbf{T}_0 ,$$

and for $i \neq 0$, using (2),

$$\mathbf{T}^* \mathbf{P}_i^* \mathbf{P}_i \mathbf{T} = (\lambda_i \mathbf{P}_i)^* (\lambda_i \mathbf{P}_i) = \bar{\lambda}_i \lambda_i \mathbf{P}_i^* \mathbf{P}_i = \mathbf{P}_i^* \mathbf{P}_i .$$

Summing these results we get

$$\mathbf{T}^* \mathbf{A} \mathbf{T} = \mathbf{A} - \mathbf{P}_0^* \mathbf{P}_0 .$$

From this relation it follows for any $\mathbf{y} \in \mathcal{R}$

$$\begin{aligned} \|\mathbf{Q} \mathbf{T} \mathbf{y}\|^2 &= (\mathbf{Q} \mathbf{T} \mathbf{y}, \mathbf{Q} \mathbf{T} \mathbf{y}) = (\mathbf{T}^* \mathbf{Q}^2 \mathbf{T} \mathbf{y}, \mathbf{y}) = (\mathbf{T}^* \mathbf{A} \mathbf{T} \mathbf{y}, \mathbf{y}) = \\ &= (\mathbf{A} \mathbf{y}, \mathbf{y}) - (\mathbf{P}_0^* \mathbf{P}_0 \mathbf{y}, \mathbf{y}) = (\mathbf{Q} \mathbf{y}, \mathbf{Q} \mathbf{y}) - (\mathbf{P}_0 \mathbf{y}, \mathbf{P}_0 \mathbf{y}) \leq \|\mathbf{Q} \mathbf{y}\|^2 . \end{aligned}$$

For any $\mathbf{x} \in \mathcal{R}$ put $\mathbf{y} = \mathbf{Q}^{-1} \mathbf{x}$, then we get from this result that

$$\|\mathbf{Q} \mathbf{T} \mathbf{Q}^{-1} \mathbf{x}\| \leq \|\mathbf{x}\| ;$$

thus $\mathbf{Q} \mathbf{T} \mathbf{Q}^{-1}$ is a contraction.

This concludes the proof of the theorem.

(Received January 19, 1959.)

REFERENCES

- [1] SZ.-NAGY BÉLA: "On uniformly bounded linear transformations in Hilbert space", *Acta Scientiarum Mathematicarum (Szeged)* **II** (1946/48), 152—157.
- [2] RIESZ F.—SZ.-NAGY B.: *Leçons d'analyse fonctionnelle*, 3. édition, Budapest—Paris, 1955.

TELJESEN FOLYTONOS OPERÁTOROKRÓL, AMELYEK ITERÁLTJAI EGYENLETESEN KORLÁTOSAK

SZÓKEFALVI-NAGY BÉLA

Kivonat

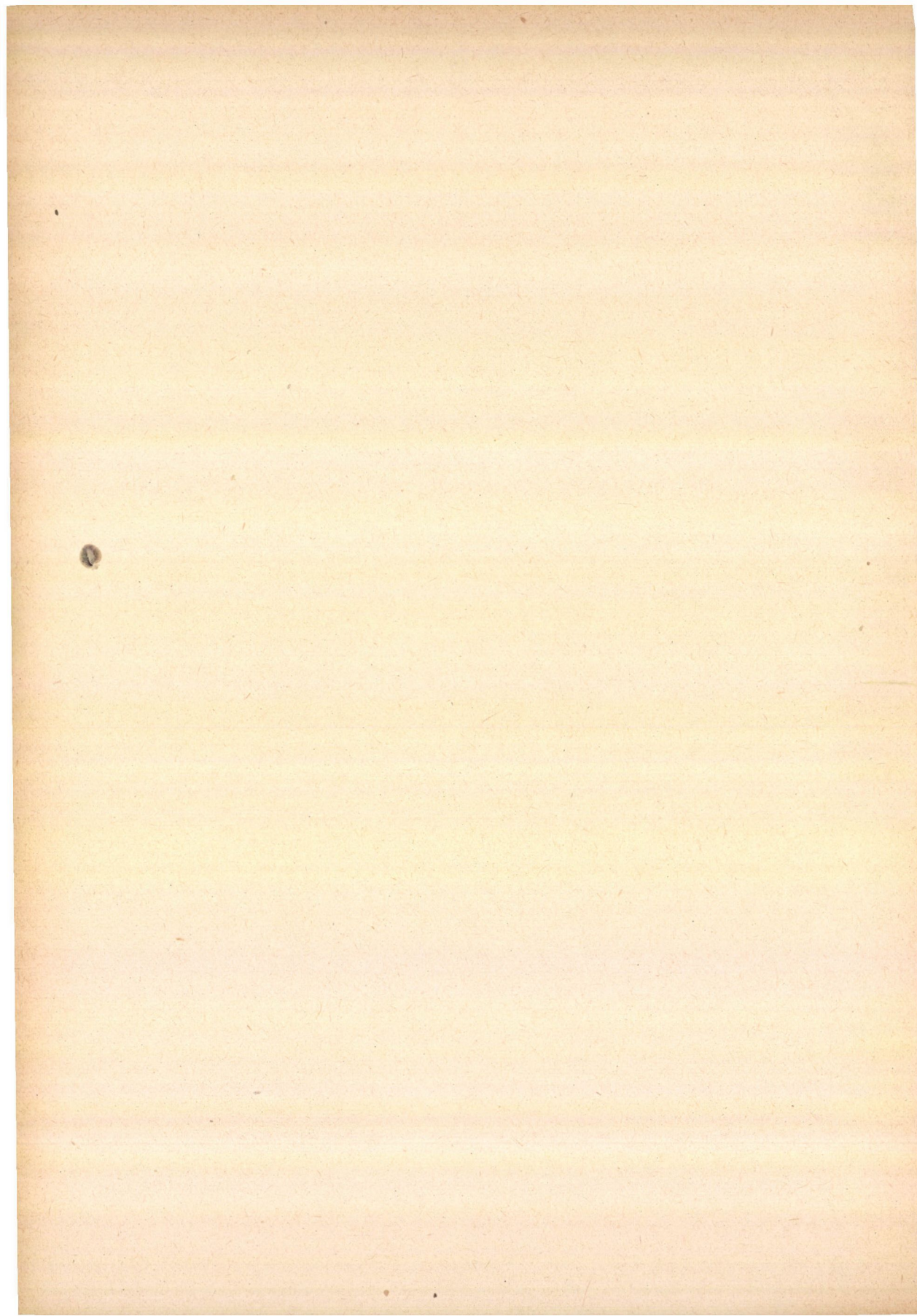
A következő tételt bizonyítjuk be: Legyen \mathbf{T} a véges vagy végtelen dimenziós komplex euklidesi tér teljesen folytonos operátora, amelyre $\|\mathbf{T}^n\| \leq M$ ($n = 1, 2, \dots$), ahol M állandó. Ekkor létezik egy olyan korlátos, és korlátos inverzű, önadjugált \mathbf{Q} operátor, amelyre a $\mathbf{T}' = \mathbf{Q} \mathbf{T} \mathbf{Q}^{-1}$ operátor kontrakció, azaz amelyre $\|\mathbf{T}'\| \leq 1$.

О ВПОЛНЕ НЕПРЕРЫВНЫХ ОПЕРАТОРАХ С РАВНОМЕРНО ОГРАНИЧЕННЫМИ ИТЕРИРОВАННЫМИ

BÉLA SZ.-NAGY

Резюме

Доказывается следующая теорема: Пусть T вполне непрерывный оператор в комплексном евклидовом пространстве конечной или бесконечной размерности, такой, что $\|T^n\| \leq M$ ($n = 1, 2, \dots$), где M постоянное. Тогда существует самосопряженный оператор Q , ограниченный и с ограниченным обратным Q^{-1} , такой, что оператор $T' = QTQ^{-1}$ является сжатием, т. е. $\|T'\| \leq 1$.



ÜBER ÄHNLICHE LINEARE TRANSFORMATIONEN IN ENDLICHDIMENSIONALEN RÄUMEN

von

L. GEHÉR

Wir nennen eine lineare Transformation T eines linearen normierten Raumes *gleichmässig beschränkt*, wenn es eine positive Konstante K derart gibt, dass $\|T^n\| \leq K$ für $n = 0, 1, 2, \dots$ gilt.¹

B. SZ.-NAGY [1] hat den folgenden Satz bewiesen:

Jede gleichmässig beschränkte vollstetige lineare Transformation T eines Hilbertschen Raumes ist ähnlich einer Kontraktion, d. h. es existiert eine umkehrbare lineare Transformation A mit $\|A T A^{-1}\| \leq 1$.

B. SZ.-NAGY stellte die Frage, ob dieser Satz nicht nur in Hilbertschen Räumen, sondern auch in allgemeinen linearen normierten Räumen (d. h. in Banachschen Räumen) gültig ist? Hier werden wir mit einer matrixtheoretischen Methode zeigen, dass dieser Satz in jedem Raum l_n^p gilt: l_n^p ($p \geq 1$) ist der lineare Raum aller n -Tupel $x = (x_1, x_2, \dots, x_n)$ von komplexen Zahlen, mit der Norm

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$$

wenn $1 \leq p < \infty$ und mit der Norm

$$\|x\|_\infty = \max_i |x_i|$$

wenn $p = \infty$.

Wir werden mit dem Beweis einer einfachen Behauptung für Matrizen anfangen.

1. Wir bezeichnen mit I_r bzw. N_r die $r \times r$ -Matrizen

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

¹ Unter einer linearen Transformation eines linearen normierten Raumes (in sich) werden wir immer eine überall definierte, beschränkte lineare Transformation verstehen. Die Norm $\|T\|$ einer linearen Transformation T wird durch $\|T\| = \sup_{\|x\|=1} \|Tx\|$ definiert.

Wir nennen eine quadratische Matrix \mathbf{J} eine Jordansche Matrix, wenn sie sich als direkte Summe endlich vieler Matrizen von der Form $\lambda \mathbf{I}_r + \mathbf{N}_r$, d. h. in der Form

$$\mathbf{J} = \sum_{i=1}^k (\lambda_i \mathbf{I}_{r_i} + \mathbf{N}_{r_i})$$

darstellen lässt. Offenbar besitzt dann \mathbf{J} $\sum_{i=1}^k r_i$ Zeilen und Spalten.

Es sei \mathbf{T} eine beliebige $n \times n$ -Matrix. Bekanntlich ist \mathbf{T} ähnlich einer und nur einer Jordanschen Matrix [2], d. h. es existiert eine umkehrbare $n \times n$ -Matrix \mathbf{A} mit

$$\mathbf{A} \mathbf{T} \mathbf{A}^{-1} = \sum_{i=1}^k (\lambda_i \mathbf{I}_{r_i} + \mathbf{N}_{r_i})$$

Wir behaupten:

Lemma. *Es existiert eine Folge $\mathbf{B}_1, \mathbf{B}_2, \dots$ von umkehrbaren $n \times n$ -Matrizen mit*

$$\mathbf{B}_m \mathbf{T} \mathbf{B}_m^{-1} \rightarrow \sum_{i=1}^k \lambda_i \mathbf{I}_{r_i} \quad \text{für} \quad m \rightarrow \infty.$$

Beweis. Es seien

$$\mathbf{C}_m = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & m & 0 & \dots & 0 \\ 0 & 0 & m^2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & m^{n-1} \end{pmatrix}$$

und $\mathbf{C}_m \mathbf{A} = \mathbf{B}_m$; beide sind umkehrbare Matrizen. Dann ist

$$\begin{aligned} \mathbf{B}_m \mathbf{T} \mathbf{B}_m^{-1} &= \mathbf{C}_m \mathbf{A} \mathbf{T} \mathbf{A}^{-1} \mathbf{C}_m^{-1} = \mathbf{C}_m \mathbf{J} \mathbf{C}_m^{-1} = \\ &= \sum_{i=1}^k \left(\lambda_i \mathbf{I}_{r_i} + \frac{1}{m} \mathbf{N}_{r_i} \right) \rightarrow \sum_{i=1}^k \lambda_i \mathbf{I}_{r_i} \end{aligned}$$

für $m \rightarrow \infty$, w. z. b. w.

2. Es sei T eine lineare Transformation von l_n^p ($p \geq 1$). (Es soll in folgendem die Matrix einer linearen Transformation von l_n^p mit demselben, aber dicken Buchstaben bezeichnet werden.) Es sei

$$\mathbf{J} = \sum_{i=1}^k (\lambda_i \mathbf{I}_{r_i} + \mathbf{N}_{r_i})$$

die Jordansche Normalform der Matrix \mathbf{T} von T , und wir setzen

$$\mathbf{D} = \sum_{i=1}^k \lambda_i \mathbf{I}_{r_i}.$$

Nach dem Lemma ist $B_m T B_m^{-1} \rightarrow D$ für $m \rightarrow \infty$. Aus der Konvergenz der Matrizen folgt aber die Konvergenz in Norm der linearen Transformationen:

$$(1) \quad \lim_{m \rightarrow \infty} \|B_m T B_m^{-1} - D\|_p = 0$$

(dies gilt auch in jedem endlichdimensionalen linearen normierten Raum)
Offenbar ist $\|D\|_p = \max_{1 \leq i \leq k} |\lambda_i|$. Nach (1) hat man dann

$$\inf_m \|B_m T B_m^{-1}\|_p \leq \inf_m \|B_m T B_m^{-1} - D\|_p + \|D\|_p = \|D\|_p = \max_{1 \leq i \leq k} |\lambda_i|.$$

Wir klassifizieren nun die Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_k$ von T folgendermassen: zur ersten Klasse gehören diejenigen Eigenwerte von T , die in der Jordanschen Normalform J von T auch in einem Block von mindestens zweiter Ordnung vorkommen, zur zweiten Klasse gehören die übrigen Eigenwerte von T . Wir können mit einer Umnummerierung immer erreichen, dass die zur ersten Klasse gehörigen Eigenwerte (falls diese Klasse nicht leer ist) gleich $\lambda_1, \lambda_2, \dots, \lambda_j$ sind.

Wir nehmen an, dass die zweite Klasse nicht leer ist und die Bedingung

$$(B) \quad \max_{1 \leq i \leq j} |\lambda_i| < \max_{j \leq i \leq k} |\lambda_i|$$

erfüllt ist. (Im Falle, dass die erste Klasse leer ist, soll $0 < \max_{1 \leq i \leq k} |\lambda_i|$ angenommen werden.)

Wir betrachten die durch

$$P(x_1, x_2, \dots, x_n) = (x_1, x_2, \dots, x_s, 0, \dots, 0)$$

und

$$Q(x_1, x_2, \dots, x_n) = (0, 0, \dots, 0, x_{s+1}, x_{s+2}, \dots, x_n)$$

definierten Projektionen P und Q ; dabei ist $s = \sum_{i=1}^j r_i$. Die linearen Transformationen D und $T_m = B_m T B_m^{-1}$ sind offenbar mit P und Q vertauschbar.

Fall 1. $1 \leq p < \infty$. Wir wählen m so gross, dass die Bedingung

$$(2) \quad \|PT_m\|_p < \|D\|_p$$

erfüllt ist. Das ist möglich, weil $\|PT_m - PD\|_p \leq \|T_m - D\|_p \rightarrow 0$ für

$m \rightarrow \infty$, und $PD = \sum_{i=1}^j \lambda_i I_{r_i}$, also nach (B)

$$\|PD\|_p = \max_{1 \leq i \leq j} |\lambda_i| < \max_{1 \leq i \leq k} |\lambda_i| = \|D\|_p$$

gilt. Da

$$QT_m = QD = DQ$$

ist, so gilt für jedes Element $x \in l_n^p$ die folgende Ungleichung :

$$\begin{aligned} \|T_m x\|_p^p &= \|PT_m Px + QT_m Qx\|_p^p = \|PT_m Px\|_p^p + \|QT_m Qx\|_p^p \leq \\ &\leq \|D\|_p^p \|Px\|_p^p + \|D\|_p^p \|Qx\|_p^p = \|D\|_p^p (\|Px\|_p^p + \|Qx\|_p^p) = \\ &= \|D\|_p^p \|x\|_p^p = \left(\max_{1 \leq i \leq k} |\lambda_i|\right)^p \|x\|_p^p \end{aligned}$$

d. h.

$$(3) \quad \|B_m T B_m^{-1}\|_p = \max_{1 \leq i \leq k} |\lambda_i| \quad \text{für} \quad m \geq m_0.$$

Fall 2: $p = \infty$. Dann ist $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$. Wir wählen m wieder so gross, dass (2) erfüllt ist. Dann ist für jedes $x \in l_n^\infty$

$$\begin{aligned} \|T_m x\|_\infty &= \|PT_m Px + QT_m Qx\|_\infty = \\ &= \max\{\|PT_m Px\|_\infty, \|QT_m Qx\|_\infty\} \leq \\ &\leq \max\{\|D\|_\infty \|Px\|_\infty, \|D\|_\infty \|Qx\|_\infty\} = \|D\|_\infty \|x\|_\infty, \end{aligned}$$

also besteht (3) auch im Falle $p = \infty$.

Damit haben wir den folgenden Satz bewiesen :

Satz 1. a) Die untere Grenze der Normen der linearen Transformationen des Raumes l_n^p ($1 \leq p \leq \infty$), die einer gegebenen linearen Transformation T von l_n^p ähnlich sind, ist $\max_{1 \leq i \leq k} |\lambda_i|$, wo $\lambda_1, \lambda_2, \dots, \lambda_k$ die Eigenwerte von T bedeuten.

b) Ist für die Eigenwerte von T auch die Bedingung (B) erfüllt, so existiert eine umkehrbare lineare Transformation B von l_n^p mit $\|BTB^{-1}\|_p = \max_{1 \leq i \leq k} |\lambda_i|$.

Ist T speziell gleichmässig beschränkt, so gilt der folgende

Satz 2. Jede gleichmässig beschränkte lineare Transformation T von l_n^p ($1 \leq p \leq \infty$) ist ähnlich einer Kontraktion von l_n^p , d. h. es existiert eine umkehrbare lineare Transformation B von l_n^p mit $\|BTB^{-1}\|_p \leq 1$.

Beweis. Offenbar ist dann $\max_{1 \leq i \leq k} |\lambda_i| \leq 1$.

Es sind zwei Fälle zu unterscheiden :

Fall 1: $\max_{1 \leq i \leq k} |\lambda_i| < 1$. Dann folgt die Behauptung aus Satz 1. a).

Fall 2: $\max_{1 \leq i \leq k} |\lambda_i| = 1$. Dann ist (B) erfüllt. Der Vollständigkeit halber

wiederholen wir den in [1] gegebenen Beweis: Ist λ ein Eigenwert von T , der auch in einem Block von J von mindestens zweiter Ordnung vorkommt, so gibt es zwei Elemente $x, y \in l_n^p$, $x \neq 0$ mit $Tx = \lambda x$ und $Ty = x + \lambda y$. Dann ist $T^m y = m \lambda^{m-1} x + \lambda^m y$. Hieraus folgt $\lim_{m \rightarrow \infty} \|T^m y\|_p = \infty$, wenn $|\lambda| = 1$ ist. Das ist aber in Widerspruch mit der gleichmässigen Beschränktheit von T .

Für $p = 2$ ist unser Satz 2. in dem erwähnten Satz von Sz.-NAGY B. [1] enthalten.

(Eingegangen 31. Januar 1959.)

LITERATURVERZEICHNIS

- [1] SZ.-NAGY B.: »Completely continuous operators with uniformly bounded iterates«, *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei*.
 [2] KELLER, O. H.: *Analytische Geometrie und lineare Algebra*, VEB Deutscher Verlag der Wissenschaften, Berlin, 1957.

A VÉGES DIMENZIÓJÚ TEREK
 HASONLÓ LINEÁRIS TRANSZFORMÁCIÓIRÓL

GEHÉR LÁSZLÓ

Kivonat

1. Tétel: Az l_n^p tér egy adott T lineáris transzformációjához hasonló transzformációi normájának alsó határa $\max_{1 \leq i \leq k} |\lambda_i|$, ahol $\lambda_1, \lambda_2, \dots, \lambda_k$ T -nek sajátértékei.

Ha (B) is teljesül, akkor van olyan megfordítható C lineáris transzformáció, amelyre $\|CTC^{-1}\| = \max_{1 \leq i \leq k} |\lambda_i|$.

2. Tétel: l_n^p minden egyenletesen korlátos transzformációja hasonló egy kontrakcióhoz.

О ПОДОБНЫХ ЛИНЕЙНЫХ ПРЕОБРАЗОВАНИЯХ
 КОНЕЧНОМЕРНЫХ ПРОСТРАНСТВ

L. GEHÉR

Резюме

Теорема 1. Нижняя грань норм преобразований, подобных с данным линейным преобразованием T пространства l_n^p , есть $\max_{1 \leq i \leq k} |\lambda_i|$, где $\lambda_1, \lambda_2, \dots, \lambda_k$ суть собственные значения T .

Если выполняется и (B), то существует такое обратимое линейное преобразование C , для которого $\|CTC^{-1}\| = \max_{1 \leq i \leq k} |\lambda_i|$.

Теорема 2. Всякое равномерно ограниченное преобразование l_n^p подобно с некоторой контракцией.

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

ÉRTESÍTÉS A II. MAGYAR MATEMATIKAI KONGRESSZUSRÓL

A Magyar Tudományos Akadémia és a Bolyai János Matematikai Társulat közösen rendezi meg Budapesten a II. Magyar Matematikai Kongresszust 1960. augusztus 24. és 31. között. A kongresszus megemlékezik Bolyai János halálának 100. évfordulójáról. A kongresszuson az alábbi szekciók fognak működni:

1. Algebra és számelmélet
2. Geometria és topológia
3. Analízis
4. Valószínűségszámítás és matematikai statisztika
5. Matematikai logika és matematikai gépek elmélete
6. A matematika alkalmazásai
7. A matematika története és oktatása.

A kongresszus iránt érdeklődők forduljanak a kongresszus szervezőbizottságához. Cím: Magyar Tudományos Akadémia Matematikai Kutató Intézete, Budapest, V., Reáltanoda utca 13—15. A szervezőbizottság az érdeklődők címére rendszeresen megküldi a kongresszusra vonatkozó tájékoztatókat.

СООБЩЕНИЕ О ВТОРОМ ВЕНГЕРСКОМ МАТЕМАТИЧЕСКОМ СЪЕЗДЕ

24—31-ого августа 1960-ого года в Будапеште состоится второй венгерский математический съезд. Съезд совместно организуют Венгерская Академия Наук и Математическое Общество имени Яноша Бояи. Съезд почтит память Яноша Бояи, в связи со столетием со дня его смерти. На съезде будут работать следующие секции:

1. алгебра и теория чисел
2. геометрия и топология
3. анализ
4. теория вероятностей и математическая статистика
5. математическая логика и теория математических машин
6. приложения математики
7. история математики и преподавание математики.

Интересующиеся могут обратиться к организационному комитету съезда. Адрес: Математический Институт Венгерской Академии Наук, Budapest, V., Reáltanoda u. 13/15. Организационный комитет будет регулярно посылать интересующимся информацию относительно съезда.

ANNOUNCEMENT OF THE II. HUNGARIAN MATHEMATICAL CONGRESS

The II. Hungarian Mathematical Congress will take place from 24th to 31st of August 1960, in Budapest, as organised jointly by the Hungarian Academy of Sciences and the János Bolyai Mathematical Society. The Congress will commemorate of János Bolyai, at the occasion of the centenary of his death.

Within the scope of the Congress the following sections will work.

1. Algebra and number theory
2. Geometry and topology
3. Analysis
4. Probability theory and mathematical statistics
5. Mathematical logic and theory of mathematical machines
6. Applications of mathematics
7. History of mathematics and mathematical education.

Those who show an interest in the Congress should apply to the Organizing Committee. Address: Mathematical Institute of the Hungarian Academy of Sciences, Budapest, V., Reáltanoda u. 13—15. The Organizing Committee will regularly supply the inquirers with further informations concerning the Congress.

COMMUNICATION SUR LE DEUXIÈME CONGRÈS MATHÉMATIQUE HONGROIS

L'Académie des Sciences de Hongrie et l'Association Mathématique János Bolyai organiseront en commun le deuxième Congrès Mathématique Hongrois qui se tiendra à Budapest du 24 au 31 Août 1960. Pendant le Congrès aura lieu la commémoration du 100-ième anniversaire de la mort de János Bolyai.

Dans le cadre du Congrès prendront part les sections suivantes :

1. Algèbre et théorie des nombres
2. Géométrie et topologie
3. Analyse mathématique
4. Calcul des probabilités et statistique mathématique
5. Logique mathématique et théorie des machines mathématiques
6. Application des mathématiques
7. Histoire et enseignement des mathématiques.

Les personnes intéressées au Congrès sont priées de s'informer auprès du Comité d'Organisation du Congrès. Adresse : Institut de Mathématique de l'Académie des Sciences de Hongrie, Budapest, V., Reáltanoda u. 13—15.

Le Comité d'organisation enverra systématiquement à l'adresse des intéressés les informations relatives à ce Congrès.

BEKANNTMACHUNG ÜBER DEN II. UNGARISCHEN MATHEMATISCHEN KONGRESS

Der II. Ungarische Mathematische Kongreß wird vom 24. bis 31. August 1960 in Budapest stattfinden. Der Kongreß wird eine gemeinsame Veranstaltung der Ungarischen Akademie der Wissenschaften und der János Bolyai Mathematischen Gesellschaft sein. Am Kongreß wird man János Bolyai, anläßlich des 100. Jahrestages seines Todes, gedenken. Im Rahmen des Kongresses werden die folgenden Sektionen tagen :

1. Algebra und Zahlentheorie
2. Geometrie und Topologie
3. Analysis
4. Wahrscheinlichkeitsrechnung und mathematische Statistik
5. Mathematische Logik und Theorie der mathematischen Maschinen
6. Anwendungen der Mathematik
7. Geschichte und Unterricht der Mathematik.

Die sich für den Kongreß Interessierenden sollen sich zum Organisationsausschuß des Kongresses wenden. Adress : Mathematisches Institut der Ungarischen Akademie der Wissenschaften, Budapest, V., Reáltanoda u. 13—15. Der Organisationsausschuß wird den Interessierten regelmäßig weitere Auskunft über dem Kongreß erteilen.

A kiadásért felel az Akadémiai Kiadó igazgatója

Műszaki felelős: Pataki Ferenc

A kézirat beérkezett: 1959. I. 7. — Példányszám: 800. — Terjedelem: 9 (A/5) ív

59/47749 — Akadémiai Nyomda, V., Gerlóczy u. 2. — Felelős vezető: Bernát György

СОДЕРЖАНИЕ

SCHNELL, E. : Об одной теореме об условном предельном распределении для одинакового распределенных независимых случайных величин	3
BRÓDY, A. : Об экономии предприятия и народного хозяйства	11
EGERVÁRY, J. : Комбинаторный метод для решения проблемы перевозки	15
VINCZE, I. : Двумерные теоремы о распределении и предельном распределении в теории вариационных рядов, II	29
ERDŐS, P.—RÉNYI A. : Центральная предельная теорема для выборках взятых из конечных множеств	49
DÉNES, J. : Число разложений перестановок в произведение наименьшего числа транспозиций и их связь с теорией графов	63
RÉNYI, A. : Замечания к теории «деревьев»	73
HANDSCOMB, D. C. : Об одной проблеме RÉNYI A.	87
Sz-NAGY, B. : О вполне непрерывных операторах с равномерно ограниченными итерированными	89
GENÉR, L. : О подобных линейных преобразованиях конечномерных пространств	95

INDEX

SCHNELL, E.: On a conditional limiting distribution theorem	3
BRÓDY, A.: Savings, computed on a price basis and savings computed on the basis of a technological matrix	11
EGERVÁRY, J.: Combinatorial method for solving the transport problem.....	15
VINCZE, I.: On some distributions and joint limiting distributions in the theory of order statistics, II.	29
FEDŐS, P.—RÉNYI, A.: On the central limit theorem for samples from a finite population	49
DÉNES, J.: The representation of a permutation as the product of a minimal number of transpositions, and its connection with the theory of graphs...	63
RÉNYI, A.: Some remarks on the theory of trees	73
HANDSCOMB, D. C.: A note on a problem of Rényi	87
SZ.-NAGY, B.: Completely continuous operators with uniformly bounded iterates	89
GEHÉR, L.: Über ähnliche lineare Transformationen in endlichdimensionalen Räumen	95

307.801

MAGYAR TUDOMÁNYOS AKADÉMIA

MATEMATIKAI
KUTATÓ INTÉZETÉNEK
KÖZLEMÉNYEI

IV. ÉVFOLYAM, 2. FÜZET

1959

★

ТРУДЫ
МАТЕМАТИЧЕСКОГО
ИНСТИТУТА

АКАДЕМИИ НАУК ВЕНГРИИ

ТОМ IV., ВЫПУСК 2.

1959

★

PUBLICATIONS
OF THE
MATHEMATICAL INSTITUTE
OF THE
HUNGARIAN ACADEMY OF SCIENCES
VOLUME IV., FASC. 2.

1959



1959

2

TARTALOMJEGYZÉK

EGERVÁRY J. †	107
ADLER Gy.: A hővezetés peremértékfeladatainak egy új típusa.....	109
ALMÁR I.—BALÁZS B.: Közelítő módszer a Hold mellett elhaladó űrrakéta pályájának megszerkesztésére	129
BÉKÉSSY A.—TÓTH K.: Megjegyzés Almár I. és Balázs B.: „Közelítő módszer a Hold mellett elhaladó űrrakéta pályájának megszerkesztése” c. cikkéhez	143
CSORDÁS L.—MENCZEL Gy.: Megjegyzések Medgyessy Pál: „Egy mechanikai függvény szintetizátor” című cikkéhez	149
RÉNYI A.: Autoklávok soros és párhuzamos kapcsolásáról és a keverés elméletéről	155
SZÜSZ P.: Feltételesen konvergens sorok átrendezéséről	167
VAS É.: A szekvenciális valószínűségi hánypados próba efficienciájáról.....	171
ÁDÁM A.: Kétpólusú elektromos hálózatokról, IV.	183
FÉNYES T.: A Mikusiński-féle operátorszámítás alkalmazása állandó együtthatós retardált differenciálegyenletek megoldására	191

RÖVID KÖZLEMÉNYEK

ADLER Gy.: Egy megjegyzés az űrhajózásról	197
---	-----

A MAGYAR TUDOMÁNYOS AKADÉMIA

MATEMATIKAI
KUTATÓ INTÉZETÉNEK
KÖZLEMÉNYEI

IV. ÉVFOLYAM, 2. FÜZET

1959

★

ТРУДЫ
МАТЕМАТИЧЕСКОГО
ИНСТИТУТА

АКАДЕМИИ НАУК ВЕНГРИИ

ТОМ IV., ВЫПУСК 2.

1959

★

PUBLICATIONS
OF THE
MATHEMATICAL INSTITUTE
OF THE
HUNGARIAN ACADEMY OF SCIENCES

VOLUME IV., FASC. 2.

1959



1959

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI KUTATÓ INTÉZETÉNEK
KÖZLEMÉNYEI

„A MAGYAR TUDOMÁNYOS AKADÉMIA ALKALMAZOTT MATEMATIKAI INTÉZETÉNEK KÖZLEMÉNYEI”

című kiadványsorozat folytatása

SZERKESZTI: RÉNYI ALFRÉD

TECHNIKAI SZERKESZTŐK: RÉVÉSZ PÁL, WEISZ KATALIN

A SZERKESZTŐSÉG CÍME: MATEMATIKAI KUTATÓ INTÉZET, BUDAPEST, V. REÁLTANODA U. 13/15.

A MATEMATIKAI KUTATÓ INTÉZET KÖZLEMÉNYEI az Intézet tudományos eredményeit tartalmazó és egyéb matematikai, valamint a matematika gyakorlati alkalmazásával kapcsolatos dolgozatokat közölnek. A folyóirat minden évfolyama 4 füzetből áll és körülbelül 30 nyomdaívet terjedelmű. A dolgozatok vagy valamelyik világnyelven jelennek meg, magyar és még egy világnyelven írt részletes kivonattal, vagy pedig magyarul, két világnyelven írt részletes kivonattal. Közlésre szánt dolgozatokat kérjük 2 gépelt példányban kivonattal ellátva a szerkesztő címére küldeni (Budapest V. Reáltanoda u. 13/15.).

A folyóirat, amelynek első évfolyama 1956-ban jelent meg, folytatása a „Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei” című kiadványsorozatnak, amelynek összesen három kötete jelent meg: I. kötet (1952), II. kötet (1953), III. kötet (1954).

A MATEMATIKAI KUTATÓ INTÉZET KÖZLEMÉNYEINEK előfizetési ára évfolyamonként belföldi címre 50,— Ft, külföldi címre 70,— Ft. (Kötetenként 5 \$) Belföldi megrendelések az Akadémiai Kiadón keresztül adhatók fel (Budapest V. Alkotmány u. 21., Magyar Nemzeti Bank egyszámlaszám: 15.915.111—46), külföldi megrendelések a Posta Központi Hírlap Iroda útján eszközölhetők (Budapest V. József nádor tér 1., Magyar Nemzeti Bank egyszámlaszám: 61.257). A folyóirat egyes füzetei 15,— Ft-os árában az Akadémiai Könyvesboltban kaphatók (Budapest V. Váci u. 22.) Cserekapcsolatok felvétele érdekében kérjük az Intézet Könyvtárához fordulni (Budapest V. Reáltanoda u. 13/15.)

ТРУДЫ
МАТЕМАТИЧЕСКОГО ИНСТИТУТА

АКАДЕМИИ НАУК ВЕНГРИИ

Продолжение издания

„A MAGYAR TUDOMÁNYOS AKADÉMIA ALKALMAZOTT MATEMATIKAI INTÉZETÉNEK KÖZLEMÉNYEI”

РЕДАКТОР: ALFRÉD RÉNYI

ТЕХНИЧЕСКИЕ РЕДАКТОРЫ: ПÁL RÉVÉSZ, KATALIN WEISZ

АДРЕС: МАТЕМАТИЧЕСКИЙ ИНСТИТУТ, БУДАПЕСТ, V. РЕАЛТАНОДА У. 13/15., ВЕНГРИЯ

В ТРУДАХ МАТЕМАТИЧЕСКОГО ИНСТИТУТА печатаются статьи, содержащие результаты научно-исследовательской работы Института, и другие математические работы, а также статьи, связанные с практическими приложениями математики. Каждый том журнала выходит в 4 выпусках и содержит приблизительно 30 печатных листов. Статьи публикуются либо на каком-нибудь мировом языке с подробным резюме на венгерском и каком-нибудь другом мировом языке, либо на венгерском языке с подробным резюме на двух мировых языках. Работы, предназначенные для опубликования, просим посылать в двух напечатанных на машинке экземплярах вместе с резюме в адрес редакции (Budapest, V. Reáltanoda u. 13/15.).

Журнал, первый том которого вышел в 1956 году, является продолжением издания «A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei» (Труды Института Прикладной Математики Академии Наук Венгрии) вышедшего в трёх томах: Том I. (1952), Том II. (1953), Том III. (1954).

Цена подписки на ТРУДЫ МАТЕМАТИЧЕСКОГО ИНСТИТУТА в Венгрии 50 форинтов, на заграничный адрес 70 форинтов (\$5 — за каждый том). Местные заказы принимает Издательство Академии Наук (Budapest V., Alkotmány u. 21., счёт Венгерского Национального Банка 15.915.111—46), заграничные заказы принимает Центральная Журнальная Контора Почты (Budapest V. József nádor tér 1., счёт Венгерского Национального Банка 61.257). По поводу отношения обмена просим обращаться к Библиотеке Института (Budapest V. Reáltanoda u. 13/15. Венгрия).

PUBLICATIONS
OF THE MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES

continuing the series

„A MAGYAR TUDOMÁNYOS AKADÉMIA ALKALMAZOTT MATEMATIKAI INTÉZETÉNEK KÖZLEMÉNYEI”

EDITOR: ALFRÉD RÉNYI

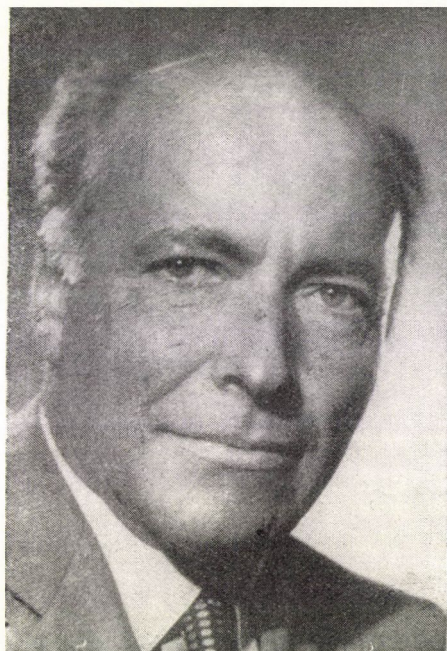
TECHNICAL EDITORS: PÁL RÉVÉSZ, KATALIN WEISZ

ADDRESS: MATHEMATICAL INSTITUTE, BUDAPEST, V., REALTANODA U. 13/15, HUNGARY

The PUBLICATIONS OF THE MATHEMATICAL INSTITUTE are publishing papers containing the results of scientific work of the Institute and other mathematical papers and papers on the practical applications of mathematics. The journal is published quarterly, 4 fasciculi are forming a volume consisting of about 30 printed lists. The papers appear either in a world-language with abstract in Hungarian and in an other world-language, or in Hungarian with abstracts in two world-languages. Papers intended for publication in the journal should be sent to the editor in 2 typewritten copies, with an abstract.

The journal, the first-volume of which appeared in 1956, continues the series «A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei» (Publications de l'Institut des Mathématiques Appliquées de l'Académie des Sciences de Hongrie) of which 3 volumes were published altogether: Volume I. (1952), Volume II. (1953), Volume III. (1954).

The price of a volume of the PUBLICATIONS OF THE MATHEMATICAL INSTITUTE is 50,— Ft to an address in Hungary and 70,— Ft (\$ 5.— The Volume) to abroad. Subscriptions can be made at the Académie Publishing House in Hungary (Budapest V. Alkotmány u. 21., single account number in the Hungarian National Bank 15.915.111—46) resp. at the Posta Központi Hírlap Iroda from abroad (Budapest V. József nádor tér 1., single-account number in the Hungarian National Bank 61.257). For establishing exchange relations please write to the Library of the Mathematical Institute (Budapest V. Reáltanoda u. 13/15., Hungary).



EGERVÁRY JENŐ
(1891—1958)

Prof. JENŐ (EUGEN) EGERVÁRY, Mitglied der Ungarischen Akademie der Wissenschaften, ist im 68 Lebensjahre verschieden. Er spielte eine bedeutende Rolle in der Begründung des Mathematischen Instituts der Ungarischen Akademie der Wissenschaften und leitete zuerst die Abteilung für Mechanik und Festigkeitslehre, später die Abteilung für Matrizentheorie und ihre Anwendungen.

Der Verstorbene hat auf zahlreichen Gebieten der Mathematik Bleibendes geschaffen. Besonders hervorzuheben sind seine geometrischen, algebraischen, matrizentheoretischen Untersuchungen, sowie seine Ergebnisse auf dem Gebiete der Differentialgleichungen und der angewandten Mathematik. Die ausführliche Aufzählung seiner Arbeiten ist in der Zeitschrift «Matematikai Lapok» (Bd. 10 (1959). S. 1—4. erschienen.

UN TYPE NOUVEAU DES PROBLÈMES AUX LIMITES DE LA CONDUCTION DE LA CHALEUR

par

GEORGES ADLER

Introduction

Il est d'usage de distinguer trois sortes de conditions aux limites traitées jusqu'ici systématiquement de l'équation de la chaleur, conditions qui peuvent être dites classiques. La première, la seconde, et la troisième sorte de conditions aux limites prescrivent respectivement sur la surface du corps la valeur de la température, la dérivée normale de la température, et la combinaison linéaire de la température et de sa dérivée normale. Ces conditions aux limites peuvent être réalisées de telle façon que le corps conducteur de la chaleur soit en contact le long de sa surface avec un réservoir de chaleur de capacité calorifique infinie (première et troisième conditions), ou bien avec une source de chaleur de capacité infinie (seconde condition). La capacité infinie signifie que la variation de la température qui se produit dans le corps conducteur de la chaleur sous l'effet du contact avec le réservoir resp. avec la source de chaleur n'exerce aucune influence sur la température du réservoir resp. sur l'intensité de la source de chaleur.

Cette conception des conditions aux limites rend évident l'introduction des conditions aux limites nouvelles correspondant au fait que la variation de la température dans le corps conducteur de la chaleur, qui est en contact avec un réservoir de chaleur, exerce une réaction sur la température de ce réservoir. Les problèmes conduisant à ce type de conditions aux limites ont déjà été traités par M. GÉZA FREUD [4] et par l'auteur [1]. Mais ils ont été examinés pour des cas linéaires extraordinairement spéciaux, et la méthode appliquée pour leurs solutions, comme nous y reviendrons dans le § 4, n'est pas convenable dans le cas d'un nombre de dimension plus grand que l'unité.

Ces nouvelles conditions correspondant à un réservoir de chaleur de capacité calorifique finie, nous les appellerons dans leur ensemble *second type* des conditions aux limites, à l'encontre du type des conditions aux limites classiques que nous appellerons dans leur ensemble de *premier type*.

Dans le § 1, nous formulerons mathématiquement le problème. Dans le § 2, nous démontrerons trois lemmes. Ces lemmes sont contenus dans un paragraphe séparé, au commencement de cette étude, d'une part afin de pouvoir s'y référer dans ce qui suit, d'autre part car nous les considérons comme intéressants en eux-mêmes aussi. Dans le § 3, nous énoncerons un principe du maximum se rapportant au système des conducteurs de la chaleur discuté dans le premier paragraphe. L'unicité de la solution de notre problème est une conséquence immédiate de ce principe du maximum. Enfin, dans le § 4, nous donnerons la solution du problème. La méthode de la solution est aussi convenable pour le calcul numérique des problèmes d'ordre pratique.

Définitions. Soient P un point de l'espace (x_1, \dots, x_m) ($m = 1, 2, 3$), t le temps, (P, t) un point de l'espace (x_1, \dots, x_m, t) . Soit C un domaine borné de l'espace (x_1, \dots, x_m) (resp. le corps conducteur de la chaleur), dont la frontière S est au cas où $m = 2$ une courbe rectifiable et au cas où $m = 3$ une surface d'aire finie [8]. Désignons par n la normale extérieure de la courbe resp. de la surface S , dont nous supposons l'existence.

Soit Σ un sous-ensemble de la frontière \mathfrak{B} du domaine \mathfrak{A} . Supposons la fonction $f(P)$ continue et bornée supérieurement sur l'ensemble $\mathfrak{A} + (\mathfrak{B} - \Sigma)$. Nous dirons alors que la fonction $f(P)$ possède une discontinuité de type \bar{A} sur la portion Σ de la frontière. Si nous remplaçons dans cette définition la borne supérieure par la borne inférieure, nous dirons que la discontinuité est de type \underline{A} . Dans le cas où une fonction possède simultanément une discontinuité de type \bar{A} et de type \underline{A} sur la portion Σ , la discontinuité sera dite de type A .

Dans toute l'étude, nous supposons que les solutions de l'équation de la chaleur possèdent des secondes dérivées continues selon les coordonnées de lieu x_i et des premières dérivées continues selon le temps t , sur l'ensemble $D = C \times (0, T)$.

§ 1. Établissement du problème

Supposons le corps borné C conducteur de la chaleur entouré par un réservoir de chaleur R (fig. 1), où il se produit au moment t une quantité $Q(t)$ de chaleur par unité de temps. Supposons de plus que la température $u(P, t)$ du corps C est une fonction du lieu et du temps satisfaisant à l'équation

$$(1) \quad \frac{\partial^2 u}{\partial x_1^2} + \dots + \frac{\partial^2 u}{\partial x_m^2} = a^2 \frac{\partial u}{\partial t} \quad (a^2 > 0)$$

de la chaleur dans le domaine C pour $t > 0$, où a est une constante matérielle. La température $f(t)$ du réservoir R ne dépend que du temps.

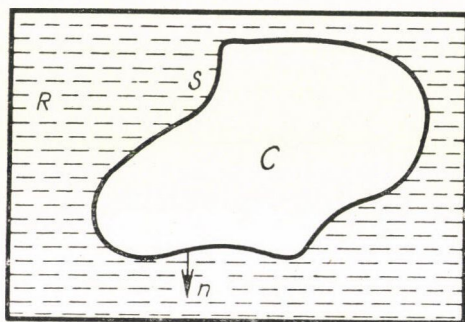


Figure 1.

On peut prescrire la condition aux limites suivante pour la fonction $u(P, t)$ à la frontière S de C , condition qui correspond au contact avec le réservoir de chaleur :

$$(a) \quad u(P, t) + \alpha \frac{\partial u(P, t)}{\partial n} = f(t) \quad (\alpha \geq 0), (t > 0, P \in S).$$

(a est une constante qui dépend de la structure de corps et du réservoir.) Les cas $a = 0$ resp. $a > 0$ correspondent à la première resp. à la troisième condition aux limites classiques. (L'analogie de la deuxième condition aux limites classique qui donnerait la valeur de la dérivée normale $\frac{\partial u}{\partial n}$, n'existe pas naturellement).

À cette condition renfermant la fonction inconnue $f(t)$ s'ajoute encore une équation de condition fournie par le bilan thermique du corps et du réservoir :

$$(b) \quad Q(t) = \beta \int_S \frac{\partial u}{\partial n} dS + \gamma f'(t) \quad (\beta, \gamma > 0).$$

(β et γ sont des constantes déterminées par les qualités matérielles et par les données géométriques.) Au premier membre de l'équation (b) figure la quantité de chaleur se produisant dans le réservoir par unité de temps, tandis que le deuxième membre est la quantité de chaleur prise pendant le même temps par le corps et par le réservoir.

En choisissant pour moment initial l'instant $t = 0$, nous pouvons donner comme conditions initiales les températures initiales du réservoir R :

$$(c_1) \quad f(0) = f_0,$$

resp. du corps C :

$$(c_2) \quad u(P, 0) = u_0(P).$$

Comme on le verra dans le § 3, les conditions (a) et (b) prises avec les conditions (c_1) et (c_2) déterminent la solution du problème.

Au cas où $a = 0$ resp. $a > 0$, nous appellerons notre problème quatrième resp. cinquième problème aux limites de la conduction de la chaleur. Ces problèmes constituent le second type des problèmes aux limites de la conduction de la chaleur. Du point de vue purement mathématique, ces problèmes aux limites de second type diffèrent de ceux du premier type par une fonction inconnue qui figure dans ces conditions nouvelles, et conformément à ce fait, ce n'est pas une seule, mais deux conditions qui appartiennent aux problèmes en question.

§ 2. Lemmes

Dans le lemme 1, nous démontrerons une généralisation du théorème généralement connu de ТЫХОНОВ [9] se rapportant au maximum des fonctions satisfaisant à l'équation de la chaleur, une généralisation pour des fonctions caloriques qui vérifient des conditions aux limites discontinues.

Lemme 1. Soit

$$D = C \times (0, T],$$

$$H = \{C \times (t = 0)\} + \{S \times [0, T]\},$$

de plus

$$\Sigma = S \times [(t = t_1) + \dots + (t = t_l)] \quad (0 \leq t_1 < \dots < t_l \leq T).$$

Supposons que la fonction $V(P, t)$ satisfasse à l'équation de la chaleur (1) dans D , soit continue dans $D + H - \Sigma$ et possède une discontinuité de type \overline{A} sur la portion Σ de la frontière de D . Soit M la borne supérieure des valeurs de la fonction prises sur $H - \Sigma$:

$$M = \sup_{(P, t) \in H - \Sigma} V(P, t).$$

Alors, la fonction $V(P, t)$ ne peut pas prendre dans D une valeur supérieure à M .

Remarque. La condition que la fonction $V(P, t)$ soit supérieurement bornée (condition qui est contenue dans celle de discontinuité de type \overline{A}) ne peut pas être supprimée. En effet, il est facile de construire une fonction calorique, dont les valeurs aux limites sont bornées, excepté un seul point de la frontière, où la fonction n'est pas définie, et les valeurs de la fonction tendent vers l'infini si l'on se rapproche de ce point de l'intérieur du domaine.

Démonstration. Supposons que la fonction prend la valeur

$$V(P^*, t^*) = M + \varepsilon \quad (\varepsilon > 0)$$

au point (P^*, t^*) de D .

Soit Σ_ϱ le voisinage de rayon ϱ de Σ dans l'espace (x_1, \dots, x_m, t) . Nous distinguons les deux cas suivants :

1°

$$N = \lim_{\varrho \rightarrow 0} \sup_{(P, t) \in \Sigma_\varrho} V(P, t) < M + \varepsilon.$$

Soit D^* la composante de l'ensemble des points satisfaisant à l'inégalité

$$V(P, t) \geq L = \frac{M + N + \varepsilon}{2} < N$$

de D , composante qui contient le point (P^*, t^*) . En vertu de la continuité de $V(P, t)$ dans $D + H - \Sigma$ et de l'inégalité

$$M + \varepsilon > L > N,$$

l'ensemble D^* est un domaine fermé et sa frontière n'a aucun point commun avec la portion Σ de la frontière de D .

Considérons l'ensemble fermé $D' = D^* \cap (t \leq t^*)$. (Au cas où $t^* = T$ on obtient $D' = D^*$.) Comme $V(P^*, t^*) = M + \varepsilon$, la portion du plan $t = t^*$ limitant le domaine D' supérieurement, portion qui appartient à la frontière de D' et qui est considérée comme un ensemble plan, contient des points intérieurs, et (P^*, t^*) en est un point intérieur. Soit H' la portion de la frontière du domaine D' , composée :

1. des points de la frontière de D' satisfaisant à l'inégalité $t < t^*$, et
2. de la frontière de l'ensemble des points vérifiant l'égalité $t = t^*$ de la frontière de D' , ensemble considéré comme un ensemble du plan $t = t^*$.

Puisque la fonction $V(P, t)$ est continue dans le domaine fermé D' et satisfait à l'équation de la chaleur (1) dans $D' - H'$, de plus, $V(P, t) = L < M + \varepsilon$ sur la portion H' de la frontière, il en résulte, selon le théo-

rème de TYKHONOFF, qu'elle ne peut pas prendre la valeur $M + \varepsilon$ au point (P^*, t^*) . Nous sommes ainsi arrivés à une contradiction.

2°

$$N = \lim_{\varrho \rightarrow 0} \sup_{(P, t) \in \Sigma_\varrho} V(P, t) \geq M + \varepsilon.$$

Soit $\pi(P, Q, t, \tau)$ la fonction calorique de la source de chaleur momentanée en activité au point (Q, τ) de l'espace (x_1, \dots, x_m, t) :

$$\pi(P, Q, t, \tau) = \begin{cases} \frac{1}{(t - \tau)^{m/2}} e^{-\frac{a^2[(x_1 - \xi_1)^2 + \dots + (x_m - \xi_m)^2]}{4(t - \tau)}}, & \text{si } t > \tau, \\ 0, & \text{si } t \leq \tau, \end{cases}$$

où $Q = (\xi_1, \dots, \xi_m)$. Cette fonction π satisfait à l'équation de la chaleur (1) dans tout l'espace (x_1, \dots, x_m, t) , excepté au point (Q, τ) de cet espace, elle est non-négative, de plus

$$(2) \quad \lim_{t \rightarrow \tau + 0} \pi(P, P, t, \tau) = +\infty.$$

Soit

$$\Pi_\Sigma(P, t) = \int_\Sigma \pi(P, Q, t, \tau) d\sigma_{(Q, \tau)}.$$

(Au cas où $m = 1$, l'intégrale \int_Σ sera remplacée par une sommation se rapportant aux points $S \times (t = t_1), \dots, S \times (t = t_l)$ qui forment Σ .) Il résulte des propriétés ci-dessus de la fonction π que la fonction Π_Σ satisfait à l'équation (1) dans tout l'espace (x_1, \dots, x_m, t) , excepté l'ensemble Σ , et

$$(3) \quad \lim_{t \rightarrow t_i + 0} \Pi_\Sigma(R, t) = +\infty \quad (R \in S), \quad (i = 1, 2, \dots, l).$$

La première partie de l'assertion est évidente, la démonstration de (3) sera effectuée plus tard.

Soit

$$\Pi_\Sigma^*(P, t) = \frac{\varepsilon}{3} \frac{\Pi_\Sigma(P, t)}{\Pi_\Sigma(P^*, t^*)}.$$

La valeur de cette fonction au point (P^*, t^*) est

$$\Pi_\Sigma^*(P^*, t^*) = \frac{\varepsilon}{3}.$$

Soit F_Σ l'ensemble des points de l'espace (x_1, \dots, x_m, t) pour lesquels

$$\Pi_\Sigma^*(P, t) > N - M.$$

En vertu de (3), l'ensemble Σ^η , qui s'obtient par une translation $\eta > 0$ suffisamment petite dans la direction de l'axe t positif, est entièrement à l'intérieur de F_Σ :

$$\Sigma^\eta = S \times [(t = t_1 + \eta) + \dots + (t = t_l + \eta)] \subset F_\Sigma.$$

Soit $\eta > 0$ aussi petit que

$$\Pi_{\Sigma}^*(P^*, t^* + \eta) < \frac{2\varepsilon}{3}$$

et

$$\Sigma^{\eta} \subset F_{\Sigma},$$

ou bien, ce qui revient au même :

$$\Sigma \subset F_{\Sigma^{(-\eta)}}.$$

À l'aide de η ainsi défini, nous construisons la fonction suivante :

$$U(P, t) = V(P, t) - \Pi_{\Sigma^{(-\eta)}}^*(P, t).$$

Cette fonction satisfait à l'équation (1) dans D , est continue dans $D + H - (\Sigma + \Sigma^{(-\eta)})$, possède une discontinuité de type \bar{A} sur les ensembles Σ et $\Sigma^{(-\eta)}$, la borne supérieure de ses valeurs prises sur la portion $H - (\Sigma + \Sigma^{(-\eta)})$ de la frontière de D est inférieure à M :

$$\sup_{(P, t) \in H - (\Sigma + \Sigma^{(-\eta)})} U(P, t) < M,$$

(car nous avons diminué la fonction V par une fonction non négative), au point (P^*, t^*)

$$U(P^*, t^*) = V(P^*, t^*) - \Pi_{\Sigma^{(-\eta)}}^*(P^*, t^*) > M + \varepsilon - \frac{2\varepsilon}{3} = M + \frac{\varepsilon}{3},$$

et finalement

$$\lim_{\varrho \rightarrow 0} \sup_{(P, t) \in \Sigma_{\varrho} + \Sigma_{\varrho}^{(-\eta)}} U(P, t) < N - (N - M) = M < M + \frac{\varepsilon}{3}.$$

Selon le cas 1° déjà démontré, cela est impossible.

Il nous reste à démontrer l'assertion (3). Pour simplifier la discussion, nous nous occuperons du cas où $m = 2$. (Au cas où $m = 1$, (3) est une conséquence immédiate de (2). Dans le cas $m = 3$, la démonstration peut se faire par analogie au cas $m = 2$.)

Vu que dans l'intégrale

$$\Pi_{\Sigma} = \int_{\Sigma} \pi d\sigma$$

la fonction à intégrer est non-négative, il suffit de nous limiter au cas où $l = 1$. Si (3) est valable pour $l = 1$, il est a fortiori valable pour $l > 1$.

Nous diminuons la valeur de l'intégrale Π_{Σ} si nous intégrons seulement sur la portion Σ' de la courbe Σ au lieu de toute la courbe Σ :

$$\Pi_{\Sigma} > \Pi_{\Sigma'},$$

où

$$\Sigma \supset \Sigma'.$$

Soit Σ' une portion connexe de Σ , qui contient le point (R, t_1) , où $R \in S$, comme extrémité. Soit g une demi-droite du plan $t = t_1$ dépliant du point (R, t_1) ,

et qui est parallèle à l'axe x_1 (fig. 2). Nous représentons la courbe Σ' sur la demi-droite g de telle façon qu'à tout point Q de Σ' corresponde un point Q' de la demi-droite g qui satisfait à la condition $\overline{Q'R} = \overline{QR}$. (Cette représentation n'est pas en général biunivoque, car différents points Q peuvent avoir le même point Q' pour image.) Soit Σ'' l'image ainsi obtenue de Σ' .

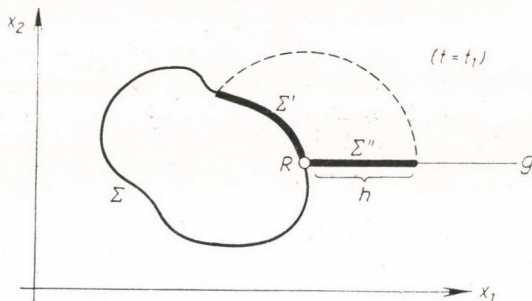


Figure 2.

Comme au cours de cette représentation, les quantités

$$t - \tau \quad (t = t_1)$$

et

$$(x_1 - \xi_1)^2 + (x_2 - \xi_2)^2 \quad ((x_1, x_2) = R)$$

ne varient pas, on obtient :

$$\pi(R, Q, t, \tau) = \pi(R, Q', t, \tau).$$

Mais, en considérant qu'au cours de la représentation, les éléments d'arc de Σ' ne sont pas plus courts que ceux correspondants de Σ'' , on a, pour cette raison

$$\Pi_{\Sigma'} \geq \Pi_{\Sigma''}.$$

Sans restriction de la généralité, soit $x_1 = x_2 = 0$ et $t_1 = 0$. Soit h la longueur de Σ'' . Alors

$$\Pi_{\Sigma''}(R, t) = \int_0^h \frac{1}{t} e^{-\frac{a^2 \xi^2}{4t}} d\xi = \frac{1}{\sqrt{t}} \int_0^{\frac{h}{\sqrt{t}}} e^{-\frac{a^2}{4} \xi^2} d\xi.$$

Étant donné que

$$\lim_{t \rightarrow 0} \Pi_{\Sigma''}(R, t) = +\infty,$$

nous avons démontré notre assertion. C. q. f. d.

Dans ce qui suit, nous utiliserons un théorème de M. L. NIRENBERG [7], dans lequel il donne une amélioration du théorème de TYKHONOFF. Dans le cas du domaine et de l'équation plus simples que nous avons examinés, ce théorème peut se formuler comme il suit :

Supposons que la fonction $V(P, t)$ possède des secondes dérivées continues selon les coordonnées x_1, \dots, x_m , et une première dérivée continue selon la variable t sur l'ensemble $C \times (0, T]$, et satisfait ici à l'équation de la chaleur (1). Si $V(P, t)$ prend le maximum de ses valeurs prises sur $C \times (0, T]$ au point (P_0, T) de $C \times (0, T]$, alors, sur $C \times (0, T]$ tout entier, $V(P, t) \equiv V(P_0, T)$.

Lemme 2. Soit P_0 un point de la frontière S du domaine borné C qui possède la propriété suivante à savoir que C contient l'intérieur d'une hypersphère Γ , dont P_0 est sur la frontière. La fonction $V(P, t)$ non identiquement constante satisfait sur $C \times (0, T]$ à l'équation (1), est continue dans $B = \{C \times (0, T]\} + (P_0, T)$ (où $P_0 \in S$), et prend le minimum de ses valeurs prises dans B au point (P_0, T) . Alors

$$\lim_{\substack{P \rightarrow P_0 \\ P \in l}} \inf \frac{V(P, T) - V(P_0, T)}{PP_0} > 0,$$

où l est une demi-droite partant de P_0 , qui forme un angle aigu avec la normale intérieure de Γ appartenant à P_0 .

Démonstration. On peut inscrire dans l'hypersphère Γ une hypersphère Γ' de rayon r_0 et de centre C_0 , dont la frontière n'a que le seul point P_0 commun avec la frontière du domaine C .

On peut supposer sans restreindre la généralité que

$$\min V(P, t) = V(P_0, T) = 0.$$

Considérons le domaine limité par les hypersphères Γ' et Γ'' , où Γ'' signifie l'hypersphère $\overline{PC_0} = \frac{r_0}{2}$. Nous désignerons ce domaine par G .

Soit $0 < T_1 < T$ et signifie $h(P, t)$ la fonction satisfaisant à l'équation (1) dans $G \times (t > T_1)$ avec les conditions aux limites et avec la condition initiale suivantes :

$$\begin{aligned} (\alpha) \quad & h(P, t) = 0 & (P, t) \in \Gamma' \times (t \geq T_1), \\ (\beta) \quad & h(P, t) = 1 & (P, t) \in \Gamma'' \times (t > T_1), \\ (\gamma) \quad & h(P, T_1) = 0 & P \in G, \end{aligned}$$

et admettant une discontinuité de type A sur $\Gamma'' \times (t = T_1)$.

Il résulte du théorème de NIRENBERG que

$$(\alpha') \quad V(P, t) > 0 \quad (P, t) \in C \times (0, T],$$

et de cette façon

$$(\beta') \quad V(P, t) \geq \varepsilon \quad (P, t) \in \Gamma'' \times [T_1, T],$$

$$(\gamma') \quad V(P, T_1) > 0 \quad P \in G.$$

Donc, en vertu du théorème de NIRENBERG, on a pour la fonction $u = V - \varepsilon h$:

$$u(P, t) = V(P, t) - \varepsilon h(P, t) \geq 0^1 \quad (P, t) \in G \times [T_1, T].$$

(En effet, supposons que contrairement à notre assertion pour un point $(P^*, t^*) \in G \times [T_1, T]$

$$u(P^*, t^*) = \mu < 0.$$

Il découle naturellement de la définition de $h(P, t)$ que $t^* \neq T_1$.

Sans restriction de la généralité, nous pouvons nous borner aux deux cas suivants :

1. $u(P, t)$ prend le minimum de ses valeurs prises dans $G \times (T_1, T]$ à l'intérieur de ce domaine, et ce minimum est $\mu = u(P^*, t^*)$. C'est impossible selon le théorème de NIRENBERG.

2. $u(P, t)$ ne prend pas le minimum indiqué ci-dessus à un point intérieur de $G \times (T_1, T]$. Alors il existe une suite (P_i, t_i) ($i = 1, 2, \dots$) des points intérieurs de $G \times (T_1, T]$, pour laquelle

$$(P_i, t_i) \rightarrow (\bar{P}, \bar{t}) \in \{G \times (t = T_1)\} + \{\Gamma' \times [T_1, T]\} + \{\Gamma'' \times [T_1, T]\},$$

et

$$u(P_i, t_i) < \mu < 0 \quad (i = 1, 2, \dots).$$

Mais c'est impossible, parce qu'au cas où

$$a) \quad (\bar{P}, \bar{t}) \in Z = \{G \times (t = T_1)\} + \{\Gamma'' \times [T_1, T]\},$$

l'existence d'une telle suite (P_i, t_i) contredit aux conditions (β) , (γ) , (β') , (γ') , en tenant compte de la continuité de $V(P, t)$ sur Z et de la validité du lemme 1 concernant la fonction $h(P, t)$; tandis qu'au cas où

$$b) \quad (\bar{P}, \bar{t}) \in \Gamma' \times [T_1, T],$$

l'existence de la suite (P_i, t_i) contredit à la condition (α) et à l'inégalité (α') .

Puisque $u(P_0, T) = 0$, pour cette raison

$$\begin{aligned} & \liminf_{\substack{P \rightarrow P_0 \\ P \in \Gamma}} \frac{u(P, T) - u(P_0, T)}{\overline{PP_0}} = \\ & = \liminf_{\substack{P \rightarrow P_0 \\ P \in \Gamma}} \frac{V(P, T) - V(P_0, T)}{\overline{PP_0}} - \varepsilon \lim_{\substack{P \rightarrow P_0 \\ P \in \Gamma}} \frac{h(P, T) - h(P_0, T)}{\overline{PP_0}} \geq 0. \end{aligned}$$

Vu qu'ici

$$\lim_{\substack{P \rightarrow P_0 \\ P \in \Gamma}} \frac{h(P, T) - h(P_0, T)}{\overline{PP_0}} = \frac{\partial h(P_0, T)}{\partial l} > 0,$$

on en tire l'assertion. C. q. f. d.

Dans un certain sens, ce théorème ne peut pas être amélioré. C'est-à-dire, en général, le théorème n'est pas valable en un point de la frontière qui est

¹⁾ Au cas où la fonction $V(P, t)$ est supposée continue sur l'ensemble $(C + P_0) \times (0, T] \supset B$, cette assertion est une conséquence immédiate du lemme 1.

un sommet du domaine. (Naturellement, en un tel point, nous ne pouvons pas parler de normale intérieure, et ainsi la direction qui forme un angle aigu avec la normale intérieure, sera tout simplement remplacée par une direction allant du sommet vers l'intérieur du domaine.) Par exemple, la fonction

$$V = \cos \frac{\pi}{2} x_1 \cos \frac{\pi}{2} x_2 e^{-\frac{\pi^2}{2a^2} t}$$

satisfaisant à l'équation de la chaleur

$$\frac{\partial^2 V}{\partial x_1^2} + \frac{\partial^2 V}{\partial x_2^2} = a^2 \frac{\partial V}{\partial t} \quad (a^2 > 0)$$

dans le domaine $-1 < x_1 < +1$, $-1 < x_2 < +1$ du plan (x_1, x_2) , qui prend son minimum en chaque point de la frontière du domaine, possède dans tous les sommets une dérivée dans une direction intérieure quelconque égale à zéro.

* *

Désignons par $E_C^{(\alpha)}(P, t)$ la fonction (dont nous supposons l'existence dans ce qui suit) satisfaisant à l'équation (1) pour $t > 0$ dans le domaine borné C limité par la surface S ,

1. au cas où $\alpha > 0$, continue dans $(C + S) \times (t \geq 0)$,
2. au cas où $\alpha = 0$, continue dans $[C \times (t \geq 0)] + [S \times (t > 0)]$ et ayant une discontinuité de type A sur la portion $S \times (t = 0)$ de la surface du domaine,

et qui satisfait à la condition initiale et à la condition aux limites

$$E_C^{(\alpha)}(P, 0) = 0 \quad (P \in C),$$

$$E_C^{(\alpha)}(P, t) + \alpha \frac{\partial E_C^{(\alpha)}(P, t)}{\partial n} = 1 \quad (P \in S, t > 0).$$

La fonction $E_C^{(\alpha)}$ ainsi définie sera appelée fonction calorique d'unité du domaine C .

Lemme 3. Soit P_0 le point de la frontière S du domaine C ayant les deux propriétés suivantes :

1. le domaine C contient l'intérieur d'une hypersphère Γ_1 , dont P_0 est situé sur la frontière;
2. le domaine C est contenu dans l'extérieur d'une hypersphère Γ_2 , dont le point P_0 est situé sur la frontière (fig. 3).

Supposons qu'existe la fonction calorique d'unité $E_C^{(0)}$ du domaine C .

Alors :

1° au cas où $t > 0$

$$\liminf_{\substack{P \rightarrow P_0 \\ P \in C}} \frac{E_C^{(0)}(P_0, t) - E_C^{(0)}(P, t)}{PP_0}$$

et

$$\limsup_{\substack{P \rightarrow P_0 \\ P \in C}} \frac{E_C^{(0)}(P_0, t) - E_C^{(0)}(P, t)}{PP_0}$$

sont finies, où l est la normale²⁾ intérieure de la surface S au point P_0 , et
 2° pour $t \rightarrow 0$, elles sont asymptotiquement égales à la fonction $\frac{a}{\sqrt{\pi}} \frac{1}{\sqrt{t}}$.

Donc, si, pour $t > 0$, $\frac{\partial E_C^{(0)}(P_0, t)}{\partial n}$ existe, alors

$$(4) \quad \frac{\partial E_C^{(0)}(P_0, t)}{\partial n} \sim \frac{a}{\sqrt{\pi}} \frac{1}{\sqrt{t}}, \quad \text{si} \quad t \rightarrow 0.$$

(n signifie la normale extérieure de la surface S au point P_0 .)

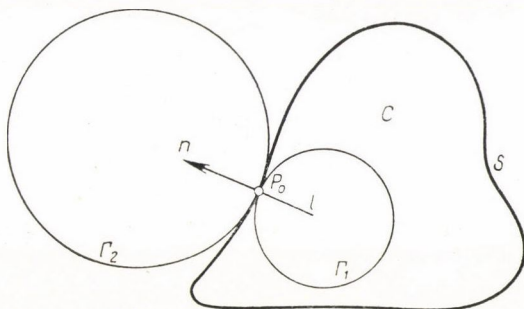


Figure 3.

Démonstration. Dans le cas $m = 1$, le domaine C est un intervalle de l'axe x_1 , ainsi l'assertion du théorème peut être vérifiée par un calcul élémentaire à partir de la forme explicite de la fonction $E_C^{(0)}(P, t)$ (voir p. e. [5]). Donc il suffit de nous limiter aux cas $m = 2, 3$.

Soient $E_{\Gamma_1}^{(0)}$ et $E_{\Gamma_2}^{(0)}$ les fonctions caloriques d'unité respectives de l'hypersphère Γ_1 et de l'extérieur de l'hypersphère Γ_2 . Soit de plus G_1 la surface de Γ_1 . Il découle de la forme explicite de Γ_2 , resp. du lemme 1, que

$$\text{resp.} \quad \left. \begin{aligned} E_{\Gamma_2}^{(0)}(P, t) &\leq 1 & (= E_C^{(0)}(P, t)), & \text{si } P \in S, \\ E_C^{(0)}(P, t) &\leq 1 & (= E_{\Gamma_1}^{(0)}(P, t)), & \text{si } P \in G_1, \end{aligned} \right\} t > 0.$$

On en tire avec l'application itérée du lemme 1, que les inégalités ci-dessus, valables sur les frontières des domaines en question, sont valables également dans les intérieurs des domaines :

$$\text{et} \quad \left. \begin{aligned} E_{\Gamma_2}^{(0)}(P, t) &\leq E_C^{(0)}(P, t), & \text{si } P \in C, \\ E_C^{(0)}(P, t) &\leq E_{\Gamma_1}^{(0)}(P, t), & \text{si } P \in \Gamma_1, \end{aligned} \right\} t > 0.$$

Donc, en considérant la relation

$$E_{\Gamma_1}^{(0)}(P_0, t) = E_C^{(0)}(P_0, t) = E_{\Gamma_2}^{(0)}(P_0, t) = 1 \quad (t > 0),$$

²⁾ L'existence de la normale au point P_0 résulte des restrictions 1. et 2.

il résulte

$$(5) \quad \frac{E_{F_1}^{(0)}(P_0, t) - E_{F_1}^{(0)}(P, t)}{\overline{PP_0}} \leq \frac{E_C^{(0)}(P_0, t) - E_C^{(0)}(P, t)}{\overline{PP_0}} \leq \frac{E_{F_2}^{(0)}(P_0, t) - E_{F_2}^{(0)}(P, t)}{\overline{PP_0}} \quad (P \in C, l).$$

Puisque dans cette inégalité, au cas où $P \rightarrow P_0$, les limites des membres gauche et droit existent (elles sont respectivement égales aux dérivées normales $\frac{\partial E_{F_1}^{(0)}}{\partial n}$ et $\frac{\partial E_{F_2}^{(0)}}{\partial n}$) l'assertion 1° en résulte.

Un calcul long mais élémentaire nous permet d'obtenir à partir des formes explicites des fonctions $E_{F_1}^{(0)}$ et $E_{F_2}^{(0)}$ (voir p. ex. [3] et [5]) :

$$\frac{\partial E_{F_1}^{(0)}(P_0, t)}{\partial n} \sim \frac{a}{\sqrt{\pi}} \frac{1}{\sqrt{t}} \quad (t \rightarrow 0)$$

et

$$\frac{\partial E_{F_2}^{(0)}(P_0, t)}{\partial n} \sim \frac{a}{\sqrt{\pi}} \frac{1}{\sqrt{t}} \quad (t \rightarrow 0).$$

D'où, en vertu de la même inégalité (5), on tire l'assertion 2°. C. q. f. d.

Remarque. Il est intéressant de noter que l'asymptotique (4) est indépendante des rayons des hypersphères F_1 et F_2 .

§ 3. Principe du maximum ; unicité de la solution

Dans ce paragraphe, nous démontrerons un principe du maximum concernant les solutions $u(P, t)$ du problème établi dans le § 1.

Supposons qu'il existe une normale en chacun des points de la frontière S du domaine borné C , de plus, qu'un point quelconque P_0 de S possède la propriété suivante : C contient l'intérieur d'une hypersphère dont P_0 est situé sur la frontière.

La fonction $u(P, t)$ satisfait à l'équation (1) de la chaleur sur l'ensemble $C \times (0, T]$, et

1. au cas $a = 0$, est continue dans

$$\{C \times [0, T]\} + \{S \times (0, T]\}$$

et possède une discontinuité de type A sur la portion $S \times (t = 0)$ de la frontière de $C \times (0, T]$,

2. au cas $a > 0$, est continue dans le domaine fermé

$$(C + S) \times [0, T].$$

Nous prescrivons les conditions aux limites suivantes pour la fonction $u(P, t)$ pour $0 < t \leq T$:

$$(6) \quad Q(t) = \beta \int_S \frac{\partial u(P, t)}{\partial n} d\sigma + \gamma f'(t) \quad (\beta, \gamma > 0) \quad \left. \vphantom{\int_S} \right\} P \in S,$$

$$(7) \quad u(P, t) + \alpha \frac{\partial u(P, t)}{\partial n} = f(t) \quad (\alpha \geq 0)$$

et nous supposons la fonction $f(t)$, déterminée par l'équation aux dérivées partielles (1) et par les conditions (6) et (7), continue dans l'intervalle $[0, T]$. (La condition pour que $f(t)$ ait une dérivée dans l'intervalle $(0, T]$ est déjà implicitement contenue dans la condition (6).)

Théorème. Si pour chaque instant $0 \leq t \leq T$

$$Q(t) \geq 0,$$

de plus

$$f(0) \geq 0, u(P, 0) \geq 0 \quad (P \in C),$$

alors dans $C \times (0, T]$:

$$u(P, t) \geq 0.$$

Démonstration. Supposons, au contraire de notre assertion, que pour un point $(P_0, t_0) \in C \times (0, T]$ l'on ait :

$$u(P_0, t_0) = \mu < 0.$$

Il résulte alors du lemme 1, que pour une valeur $0 < t_1 \leq t_0$ et pour un point $P_1 \in S$

$$u(P_1, t_1) = \mu.$$

Au cas où $a = 0$ il résulte de la continuité de $f(t)$ et de la condition (7) tandis que pour le cas où $a > 0$, il résulte de la continuité de $u(P, t)$, l'existence d'un t^* ($0 < t^* \leq t_1$), pour lequel on peut encore trouver un point $P^* \in S$ tel que

$$(8) \quad u(P^*, t^*) = \mu,$$

mais

$$(9) \quad \begin{cases} u(P, t) > \mu, \\ \text{si} \quad (P, t) \in \{C \times [0, t^*)\} + \{S \times (0, t^*)\}. \end{cases}$$

Nous montrerons que c'est impossible.

De la définition de t^* et du lemme 2, il vient

$$(10) \quad \frac{\partial u(P^*, t^*)}{\partial n} < 0.$$

1° Soit $a = 0$. De la condition (7) et de la définition de t^* , il résulte que

$$(11) \quad f'(t^*) \leq 0.$$

Nous écrirons (6) pour $t = t^*$:

$$(12) \quad Q(t^*) = \beta \int_S \frac{\partial u(P, t^*)}{\partial n} d\sigma + \gamma f'(t^*).$$

En vertu de la condition (7) :

$$u(P, t^*) = u(P^*, t^*) \quad (P \in S).$$

C'est pourquoi l'inégalité (10) est aussi valable outre P^* pour un point quelconque $P \in S$:

$$(10') \quad \frac{\partial u(P, t^*)}{\partial n} < 0 \quad (P \in S).$$

Ainsi le second membre de l'équation (12) est négatif d'après (10') et (11), tandis que son premier membre est non-négatif selon les conditions. Ainsi nous sommes arrivés à une contradiction.

2° Soit $\alpha > 0$. Nous écrivons (7) pour le point (P^*, t^*) :

$$u(P^*, t^*) + \alpha \frac{\partial u(P^*, t^*)}{\partial n} = f(t^*).$$

Il suit de (8) et de (10) que

$$f(t^*) < \mu.$$

Étant donné que $f(0) \geq 0$, on en tire l'existence d'une valeur $0 < t_2 < t^*$ pour laquelle

$$(13) \quad f(t_2) = \mu, \quad f'(t_2) \leq 0.$$

Écrivons (6) pour $t = t_2$, en utilisant la condition (7) :

$$Q(t_2) = \frac{\beta}{\alpha} \int_S [f(t_2) - u(P, t_2)] d\sigma + \gamma f'(t_2).$$

Selon cette équation il résulte de (13) et de $Q(t_2) \geq 0$ que pour certains points $P \in S$ $u(P, t_2) \leq \mu$, ce qui contredit (9) à cause de $t_2 < t^*$. C. q. f. d.

Si dans ce théorème nous supposons les inégalités contraires :

$$Q(t) \leq 0 \quad (0 \leq t \leq T)$$

et

$$f(0) \leq 0, \quad u(P, 0) \leq 0 \quad (P \in C),$$

de plus que la fonction $u(P, t)$ possède une discontinuité de type \bar{A} au lieu de celle de type A , alors l'inégalité figurant dans l'assertion du théorème sera également remplacée par l'inégalité de sens contraire :

$$u(P, t) \leq 0 \quad ((P, t) \in C \times (0, T]).$$

Corollaire. Si de plus, en conservant les conditions de continuité relatives aux fonctions $u(P, t)$ et $f(t)$ dans le théorème ci-dessus, nous supposons, au cas où $\alpha = 0$, que la fonction $u(P, t)$ a une discontinuité de type A sur la portion $S \times (t = 0)$ de la frontière, alors l'unicité du système de solution $\{u(P, t), f(t)\}$ résulte du principe du maximum ci-dessus de la manière bien connue :

En effet, s'il existait deux systèmes de solution, notamment $\{u_1, f_1\}$ et $\{u_2, f_2\}$, alors le système $\{u = u_1 - u_2, f = f_1 - f_2\}$ satisferait aux conditions initiales et aux conditions aux limites homogènes, c'est pourquoi nous aurions en même temps

$$u \geq 0, \quad u \leq 0,$$

d'où l'on obtiendrait $u \equiv 0$, et suivant la condition (7), $f \equiv 0$.

Plus généralement, il résulte également du principe du maximum démontré ci-dessus que le système de solution est une fonction continue des fonctions $Q(t)$, $u_0(P)$ et de la valeur de f_0 .

§ 4. Solution du problème

Cherchons la solution de l'équation (1) avec les conditions aux limites et initiales (a) , (b) , (c_1) et (c_2) (voir § 1). En considérant la linéarité de l'équation et celle des conditions, nous cherchons la solution sous la forme

$$u(P, t) = u_1(P, t) + u_2(P, t),$$

où u_1 satisfait aux conditions

$$u_1(P, 0) = u_0(P) \quad (P \in C),$$

$$u_1(P, t) + \alpha \frac{\partial u_1(P, t)}{\partial n} = 0 \quad (t > 0, P \in S),$$

et u_2 aux conditions

$$u_2(P, 0) = 0 \quad (P \in C),$$

$$u_2(P, t) + \alpha \frac{\partial u_2(P, t)}{\partial n} = f(t) \quad (t > 0, P \in S),$$

$$Q^*(t) = Q(t) - \beta \int_S \frac{\partial u_1}{\partial n} d\sigma = \beta \int_S \frac{\partial u_2}{\partial n} d\sigma + \gamma f'(t).$$

Vu que la détermination de $u_1(P, t)$ conduit à la solution du premier resp. du troisième problème aux limites classiques au cas où $\alpha = 0$ resp. où $\alpha > 0$, nous supposons la fonction $u_1(P, t)$ déjà connue; il nous reste donc à déterminer la fonction $u_2(P, t)$. Par suite, sans restriction de la généralité, nous pouvons nous borner au cas

$$(c_2^*) \quad u_0(P) \equiv 0.$$

* *

Le principe de la méthode de solution de M. FREUD ([4] et [1]) consiste à chercher la fonction $u(P, t)$ sous la forme

$$u(P, t) = \int_0^t \int_S W(R, \tau) \pi(P, R, t, \tau) d\sigma_R d\tau,$$

où $\pi(P, R, t, \tau)$ est la fonction calorique de la source ponctuelle de chaleur, dont nous nous sommes déjà servi dans le lemme 1 du § 2. C'est-à-dire, nous dérivons la température $u(P, t)$ à partir d'une répartition continue des sources de chaleur en activité sur la frontière du corps avec une intensité $W(P, t)$. De cette façon, nous avons ramené notre problème à la détermination de la fonction $W(P, t)$ qui est seulement la fonction des points de la surface S et du temps. Dans le cas linéaire ($m = 1$) la fonction $W(P, t)$ se réduit seulement à une fonction du temps, ainsi cette méthode simplifie de beaucoup la solution de notre problème. Par contre au cas où $m > 1$, il ne serait pas convenable

d'appliquer cette méthode. C'est pourquoi nous résolvons le problème d'une autre manière, à l'aide du principe de DUHAMEL [2]. Cette méthode nouvelle nous permet, dans le cas d'un nombre de dimensions quelconque, de ramener notre problème à la détermination d'une fonction unique ne dépendant que du temps.

**

Le principe de DUHAMEL affirme qu'en supposant la fonction calorique d'unité $E_C^{(\omega)}(P, t)$ du domaine C connue (voir § 2), la solution de l'équation différentielle (1) déterminée par les conditions

$$u(P, 0) = 0, \quad (P \in C)$$

$$u(P, t) + \alpha \frac{\partial u(P, t)}{\partial n} = f(t) \quad (t > 0, P \in S)$$

peut être obtenue sous la forme

$$(14) \quad u(P, t) = \int_0^t f'(\tau) E_C^{(\omega)}(P, t - \tau) d\tau + f(0) E_C^{(\omega)}(P, t),$$

sous certaines conditions de régularité concernant la fonction $f(t)$. (Nous reviendrons à ces conditions.)

En envisageant les conditions (a) et (c_2^*) , nous cherchons la solution de notre problème sous la forme (14). De cette façon, après avoir utilisé les conditions (a), (c_1) et (c_2^*) , il nous reste à déterminer la fonction unique $f(t)$ dépendant seulement du temps, ce qui est possible à l'aide de la condition (b) encore non utilisée.

En substituant (14) à la condition (b), on obtient :

$$(15) \quad \begin{aligned} Q(t) &= \beta \int_S \left[\int_0^t f'(\tau) \frac{\partial E_C^{(\omega)}(P, t - \tau)}{\partial n} d\tau + f_0 \frac{\partial E_C^{(\omega)}(P, t)}{\partial n} \right] d\sigma + \gamma f'(t) = \\ &= \beta \int_0^t f'(\tau) K^{(\omega)}(t - \tau) d\tau + \beta f_0 K^{(\omega)}(t) + \gamma f'(t), \end{aligned}$$

où

$$K^{(\omega)}(t) = \int_S \frac{\partial E_C^{(\omega)}(P, t)}{\partial n} d\sigma.$$

En introduisant les notations

$$\varphi(t) = f'(t),$$

$$\lambda = \frac{\beta}{\gamma},$$

$$M^{(\omega)}(t) = \frac{1}{\gamma} Q(t) - \frac{\beta}{\gamma} f_0 K^{(\omega)}(t),$$

notre équation prend la forme suivante :

$$(16) \quad \varphi(t) + \lambda \int_0^t \varphi(\tau) K^{(\alpha)}(t - \tau) d\tau = M^{(\alpha)}(t) .$$

C'est une équation intégrale de seconde espèce du type de Volterra. Le noyau $K^{(\alpha)}(t - \tau)$ de cette équation est borné pour $\alpha > 0$. Au cas où $\alpha = 0$, supposons que chaque point de la surface S possède la propriété 2. du lemme 3 du § 2 ; supposons de plus que les longueurs des rayons des sphères Γ_2 figurant dans ce lemme ont une borne inférieure positive. Alors, comme on peut le voir immédiatement par la démonstration du lemme 3, l'estimation $\frac{\partial E_C^{(0)}}{\partial n} < \frac{a + \varepsilon}{\sqrt{\pi} \sqrt{t}}$ ($\varepsilon > 0$) au cas où $t \rightarrow 0$ est valable uniformément pour chaque point de la surface S , et de cette manière, le noyau $K^{(0)}$ est de singularité faible. (Il résulte du lemme 3 que dans le cas où la surface S est suffisamment régulière, la singularité de $K^0(t)$ est justement de caractère $\frac{1}{\sqrt{t}}$.) Si nous supposons la fonction $Q(t)$ absolument intégrable dans l'intervalle $[0, T]$, alors le second membre $M^{(\alpha)}(t)$ de l'équation intégrale est aussi absolument intégrable dans l'intervalle $[0, T]$; ainsi, selon la théorie des équations intégrales [6], la solution $\varphi(t)$ existe dans l'intervalle $[0, T]$ pour une valeur quelconque de λ .

De l'équation (16) il découle que les singularités éventuelles de $\varphi(t)$ sont conformes à celles de $M^{(\alpha)}(t)$. Ainsi, si $Q(t)$ est absolument intégrable, alors la fonction $u(P, t)$ peut se mettre sous la forme (14). De plus, si nous supposons $Q(t)$ bornée, excepté le seul point $t = 0$, où elle peut avoir une singularité de caractère $\frac{1}{\sqrt{t}}$, alors dans l'expression (14) la dérivation normale peut s'exécuter dans la fonction à intégrer et l'ordre des intégrales dans l'expression (15) peut être interverti, comme il résulte de la convergence uniforme des intégrales en question.

Donc nos résultats peuvent être résumés dans le théorème suivant d'existence :

Théorème. *Supposons que la fonction calorique d'unité $E_C^{(\alpha)}(P, t)$ du domaine C existe, de même que $\frac{\partial E_C^{(\alpha)}}{\partial n}$ existe et soit intégrable sur la surface S pour $t > 0$. Au cas où $\alpha = 0$ chaque point de la surface S possède la propriété 2. du lemme 3 du § 2, et les longueurs des rayons des sphères Γ_2 figurant dans ce lemme possèdent une limite inférieure positive. Soit la fonction $Q(t)$ intégrable et bornée excepté au seul point $t = 0$, où elle peut avoir une singularité de caractère $\frac{1}{\sqrt{t}}$. Alors la solution de l'équation (1) existe dans les conditions (a), (b), (c₁) et (c₂^{*}) et elle peut s'écrire sous la forme (14) ; la détermination de $f(t)$ est possible au moyen de l'équation (16).*

(Reçu le 4 Mai 1959.)

BIBLIOGRAPHIE

- [1] ADLER Gy.: „Hővezetési és diffúziós feladatok összetett peremfeltételekkel, II.” *MTA Matematikai Kutató Intézetének Közleményei*, **1** (1956) 167–183.
- [2] COURANT, R.—HILBERT, D.: *Methoden der Mathematischen Physik*, II Verlag Springer, Berlin (1937) 183–184.
- [3] ДИТКИН, В. А.—КУЗНЕЦОВ, П. И.: *Справочник по операционному исчислению*. Гос. Изд. Техн.—Теор. Лит., Москва—Ленинград (1951) 88–90.
- [4] FREUD G.: „Hővezetési és diffúziós feladatok összetett peremfeltételekkel, I.” *MTA Alkalmazott Matematikai Intézetének Közleményei* **3** (1955) 369–394.
- [5] ЛЫКОВ, А. В.: *Теория теплопроводности*. Гос. Изд. Техн.—Теор. Лит., Москва (1952) 77–81, 95–98, 105–110.
- [6] МИХЛИН, С. Г.: *Интегральные Уравнения*. Гос. Изд. Техн.—Теор. Лит., Москва—Ленинград (1949) 23–27.
- [7] NIRENBERG, L.: „A strong maximum principle for parabolic equations.” *Comm. on Pure and Applied Mathematics* **6** (1953) 167–177.
- [8] DE LA VALLÉE POUSSIN, J. CH.: *Cours d'Analyse Infinitésimale*, I. Gauthier-Villars, Paris (1923) 359–363.
- [9] ТИХОНОВ, А. Н.—САМАРСКИЙ, А. А.: *Уравнения математической физики*. Гос. Изд. Техн.—Теор. Лит., Москва—Ленинград (1951) 184–187.

A HŐVEZETÉS PEREMÉRTÉKFELADATAINAK EGY ÚJ TÍPUSA

ADLER Gy.

Kivonat

A hővezetés három klasszikus peremfeltétele rendre a hőmérsékletet, a hőmérséklet normális menti deriváltját, illetve a hőmérséklet és a hőmérséklet normális menti deriváltjának lineáris kombinációját írja elő a test peremén. Ezek a peremfeltételek fizikailag akkor teljesülnek, ha a hővezető test környezetének állapota a hővezető testben lejátszódó folyamattól nem függ. Ezzel szemben, ha a testben végbemenő hővezetési folyamat a környezet állapotára visszahat, akkor a folyamat leírására új peremfeltételek válnak szükségessé. Pílen peremfeltételek egy típusával foglalkozik a szerző (lásd (a) és (b) feltételek). Bebizonyítja ezen új típusú probléma megoldásának unicitását és egzisztenciáját. Az egzisztencia-bizonyítás módszere konstruktív és gyakorlati feladatok megoldására alkalmas. (A szerző a probléma megoldását a Duhamel-elv segítségével másodfajú Volterra-típusú integrálegyenlet megoldására vezeti vissza.)

A második paragrafusban bebizonyított három lemma közül kiemeljük az 1. lemmát, melyben a szerző a hővezetés egyenletének elegettevő függvények maximumára vonatkozó közismert ТИХОНОВ-tétel [9] nem-folytonos peremfeltételeket kielégítő hőmérsékleti függvényekre vonatkozó kiterjesztését bizonyítja be.

ОДИН НОВЫЙ ТИП ГРАНИЧНЫХ ЗАДАЧ ТЕПЛОПРОВОДНОСТИ

Gy. ADLER

Резюме

Три классических граничных условий теплопроводности предписывают температуру, производную температуры вдоль нормали или их линейную комбинацию на поверхности тела. Эти граничные условия физически выполняются, если состояние окрестности теплопроводного тела не зави-

сит от процесса, протекающего в этом теле. Если же процесс теплопроводности, протекающий в теле, оказывает влияние на состояние окрестности, то для описания процесса требуются новые граничные условия. Автор занимается одним типом таких условий (смотри условия (a) и (b)). Доказывает существование и единственность решения этой новой проблемы. Доказательство существования конструктивно и годится для решения практических задач. (Автор с помощью принципа Дунамел-а сводит решение проблемы к решению интегрального уравнения Вольтерра второго рода).

Среди трёх лемм, доказанных во втором параграфе, отметим лемму 1, в которой автор доказывает распространение общеизвестной теоремы Тихонова [9] относительно максимума функций, удовлетворяющих уравнению теплопроводности, на температурные функции, удовлетворяющие непрерывным граничным условиям.

APPROXIMATE METHOD OF PLOTTING THE ORBIT OF A SPACE ROCKET PASSING NEAR THE MOON

by

IVÁN ALMÁR¹ and BÉLA BALÁZS¹

I. Introduction

Leaders of Moonwatch and radio tracking stations of smaller countries may frequently have to face the problem in the future: how to select best suitable times, directions and means for the observation of some space rocket just launched, when only sparse information concerning its positions reported in newspapers are available. To this purpose they will need a quick method for determining the orbit of the rocket, at least approximately. The present paper, starting from IEGOROV's ideas [1] suggests a method which lends itself to plotting the whole orbit of a space rocket the type launched on the 2nd January 1959 on the basis of minimum information. Quick applicability of the method is enhanced by nomograms.

We make the following basic assumptions:

1. The space rocket starting from beyond the atmosphere travels in free flight;

2. The plane of its orbit makes a small angle with the ecliptic and at the time of the closest approach the Moon is passing through the ecliptic;

3. The Moon describes a circular orbit around the Earth. Whereas assumption 1. with rockets flying without remote control is fulfilled from the very beginning, the soundness of 2. has to be examined in each particular case separately on the basis of the initial elements.² Assumption 3. does not imply any significant restriction. The motion of a lunar probe makes part of the problem of n bodies in celestial mechanics considered insolvable in general manner. Instead of giving the equations of motion of a space vehicle travelling in the Earth-Moon system (problème restreinte), IEGOROV [1] reduces the task to several one-centrum problems. He assumes that the motion of the third body considered to be of a negligibly small mass is influenced inside an environment (sphere) of 66 000 km in radius around the Moon only by the Moon, and outside this sphere only by the Earth.

The sphere of the Moon is defined in [2] as follows: A third body is inside the Moon sphere whenever the ratio of the perturbing force coming from the Earth and affecting the selenocentric motion, to the force of attraction of the Moon is smaller than the quotient of the force perturbing the geocentric motion (i. e. coming from the Moon) to the force of attraction of

¹ Konkoly Observatory of the Hungarian Academy of Sciences.

² The method with some slight modifications may presumably find universal application: to the necessity and possibility of its generalization we will come back later.

the Earth. A similar definition may be given to the sphere of the Earth relatively to the Sun and in general, to the sphere of a body having a mass m relatively to another body having a mass M . ($M > m$). The radius of the sphere may generally be computed as

$$(1) \quad r = A \left(\frac{m}{M} \right)^{2/5}$$

where A denotes the distance between the two bodies [2]. The sphere of the Earth relatively to the Sun has, according to (1) a radius of 930 000 km.

The voyage of a space rocket passing near the Moon may be divided the following phases:

- 1) Flight from the starting point to the boundary of the Moon's sphere (a conic section in geocentric co-ordinates);
- 2) Travel in the Moon's sphere (a conic section in selenocentric co-ordinates);
- 3) Flight beyond the Moon's sphere to the boundary of the Earth's sphere (a conic section in heliocentric co-ordinates).

II. Methods Used for Plotting the Orbit

§ 1. Conversion of Given Positions to Polar Co-ordinates

The positions reported immediately after launching mostly give information on the space rocket having reached a height h at an instant t over some point defined through geographic co-ordinates (φ , λ). We will assume that the rocket travels in the ecliptic and our task is to change over to polar co-ordinates. One of the co-ordinates is simply

$$r = h + R_{Earth}$$

whereas the projection on the Equator of the polar angle computed from the Vernal Equinox will be obtained by increasing the sidereal time S with λ , when S denotes the sidereal time corresponding to the instant t expressed in U T.³ For the local sidereal time $S + \lambda$ and the angle ϑ the relationship of the spherical trigonometry

$$(2) \quad \sin \vartheta = \frac{\sin (S + \lambda)}{\sqrt{1 - \sin^2 i \cos^2 (S + \lambda)}}$$

holds true, where $i = 23^\circ 27'$. The relationship of $S + \lambda$ and $\vartheta - (S + \lambda)$ may be read in Fig. 1.

³ East longitudes are taken with positive, West longitudes with negative signs.

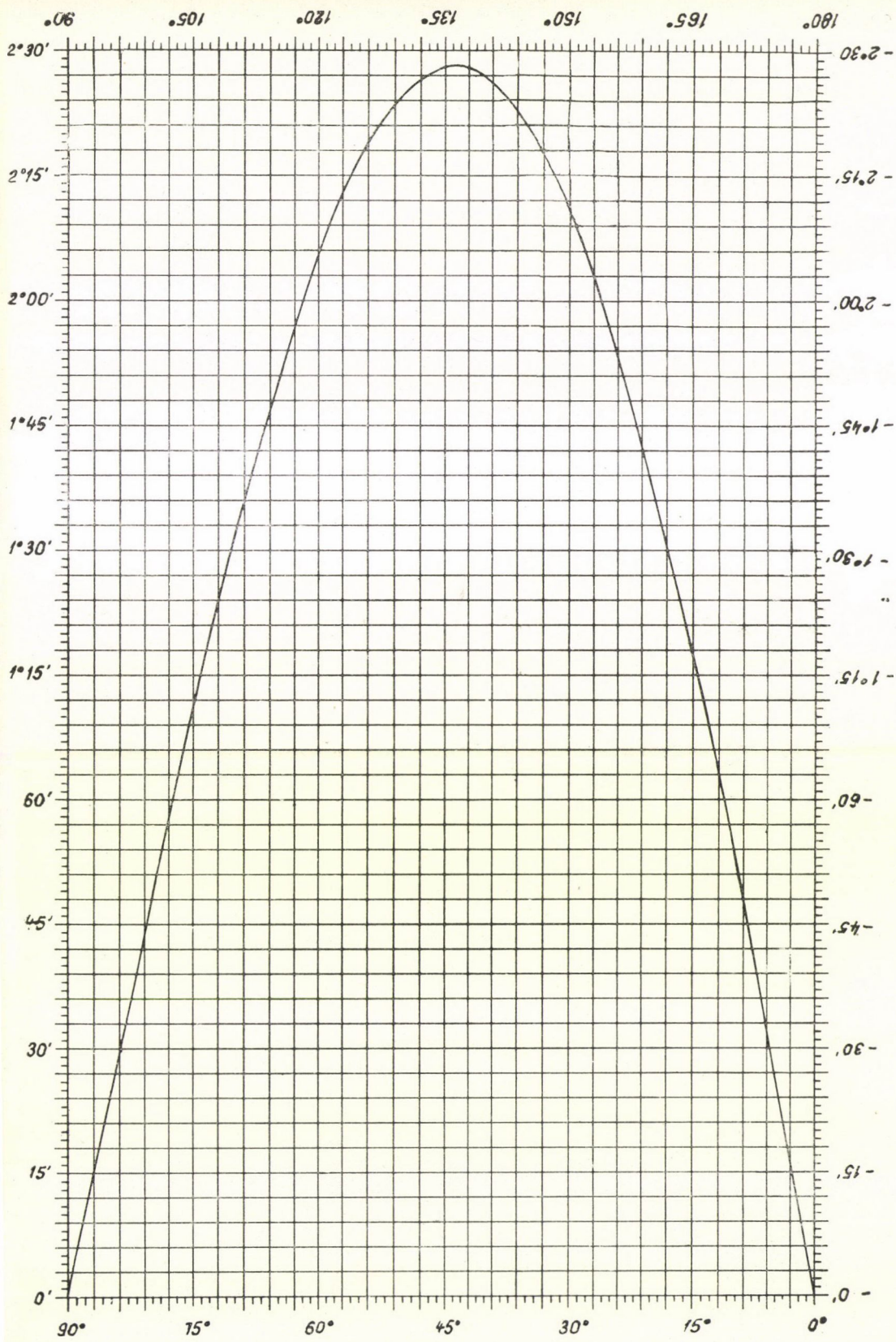


Figura 1.

§ 2. Construction of Hyperbolic Orbit from 5 Nearby Points

Method A. In case we can determine the polar referring to the hyperbola of the known focus F , the direction of the imaginary axis is determined. Construction of the polar (p) may proceed from the theorem of the complete quadrilateral, therefore, starting from the five points given (1, 2, 3, 4, 5) according to PASCAL's theorem we may locate the points 6 and 7 of the hyperbola so that the straight lines (1, 6) and (5, 7) fit to F . Then the complete quadrilateral 1, 5, 6, 7 determines the polar p of point F . The major axis

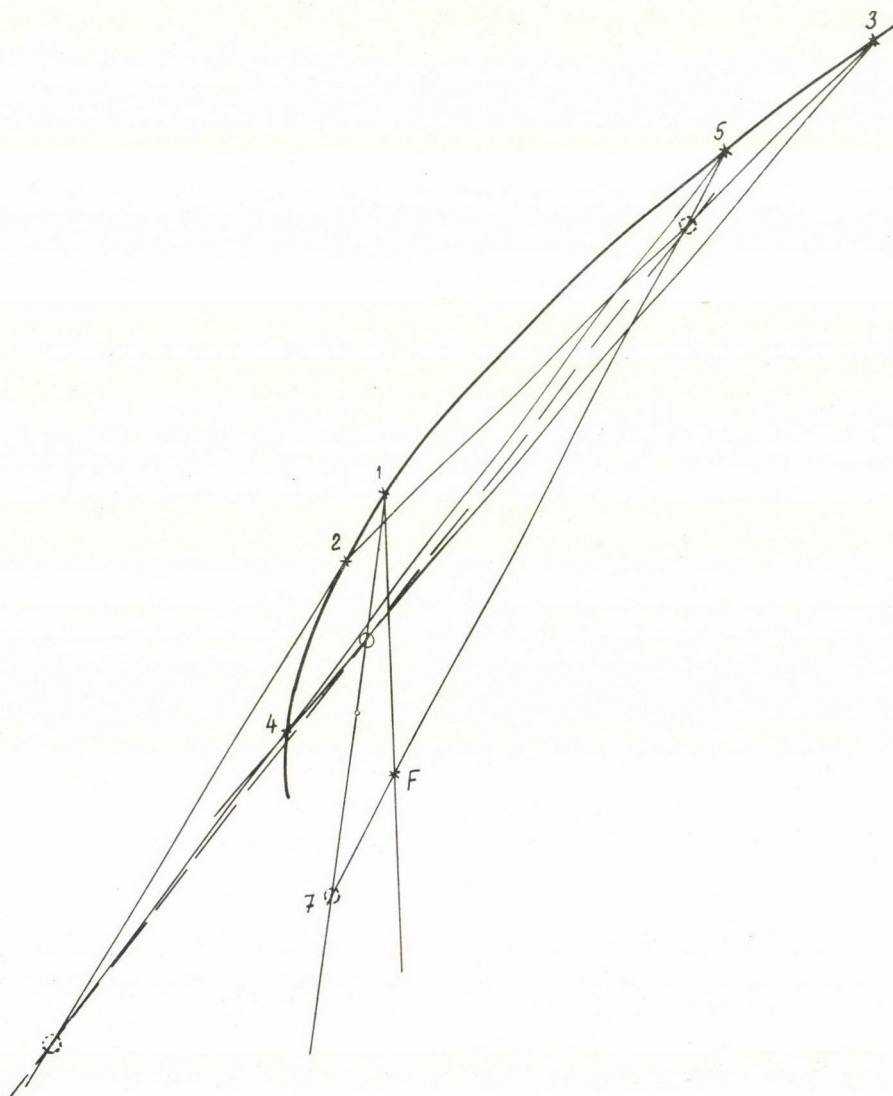


Figure 2a.

of the hyperbola is perpendicular to the polar, passes through the focus and cuts off the pericenter of the hyperbola (with the STEINERIAN construction).

Method B. From five given points (1, 2, 3, 4, 5) further two points (6, 7) on the conic may be found which, together with the points given, suffice to determine two pairs of parallel chords. For instance a secant drawn parallelly to (1, 2) from 5 cuts off point 6 and the parallel to (1, 6) drawn from 3 cuts off point 7. The mid-points of the parallel chords lie on diameters conjugate to the respective direction and thus their point of intersection gives the center of the conic. When focus and center have been found, the rest of the parameters may readily be obtained.



Figure 2b.

§ 3. Construction of Hyperbola from Remote Points

In case the method described above cannot be applied because the curvature of the hyperbolic branch is too slight, a straight line passing through the basic points which gives the minimum of scattering should be interpreted as a tangent drawn from an intermediate point and the known radius vector r_1 of the point should be reflected on this line. The mirror image indicates the direction of the second focus. Its distance, i. e. the length of the second radius vector is for a hyperbola evidently $r_1 + 2a$, where a may be determined by the process described in §. 4, by taking the average velocity of the selected straight line section as $|v|$.

§ 4. Determination of the Major Axis of the Orbit

We will use the energy integral known from celestial mechanics [3]

$$(3) \quad v^2 = fm \left(\frac{2}{r} - \frac{1}{a} \right)$$

where v denotes the velocity of the point,
 r the length of its radius vector,
 a the semi-major axis of the orbit,
 m the mass of the central celestial body, and
 f the gravitation constant.

When $|v|$ and r are known for any point of the track, a may be obtained either from Formula (3) or from one of the Nomograms 2, 3, 4 (Fig. 2 gives the nomogram corresponding to the one-centrum problem referring to the Earth, Fig. 3 that referring to the Moon, and Fig. 4 that referring to the Sun. In the case of the Moon, the relationship is needed only at the sphere boundary where $\varrho = 66\,000$ km). When a is known the construction described in §. 3 may be completed.

§ 5. Duration of Flight between Two Given Points of the Orbit

For elliptical orbits the KEPLER equation [4]

$$(4) \quad \sqrt{\frac{fm}{a^3}} (t - t_0) = u - e \sin u$$

in case of known m , a and e gives the relationship between the eccentric anomaly (u) of the selected point P on the orbit and the moment t of arrival, where the time is counted from the instant of the pericentrum transit (t_0). By projecting point P on the auxiliary circle of the ellipse parallelly to the minor axis, the distance between the point P' obtained and the major axis will be $a \cdot \sin u$, wherefrom u may be obtained.

For a hyperbola, the KEPLER equation takes the form [4]

$$(5) \quad \sqrt{\frac{fm}{a^3}} (t - t_0) = e \operatorname{sh} u - u$$

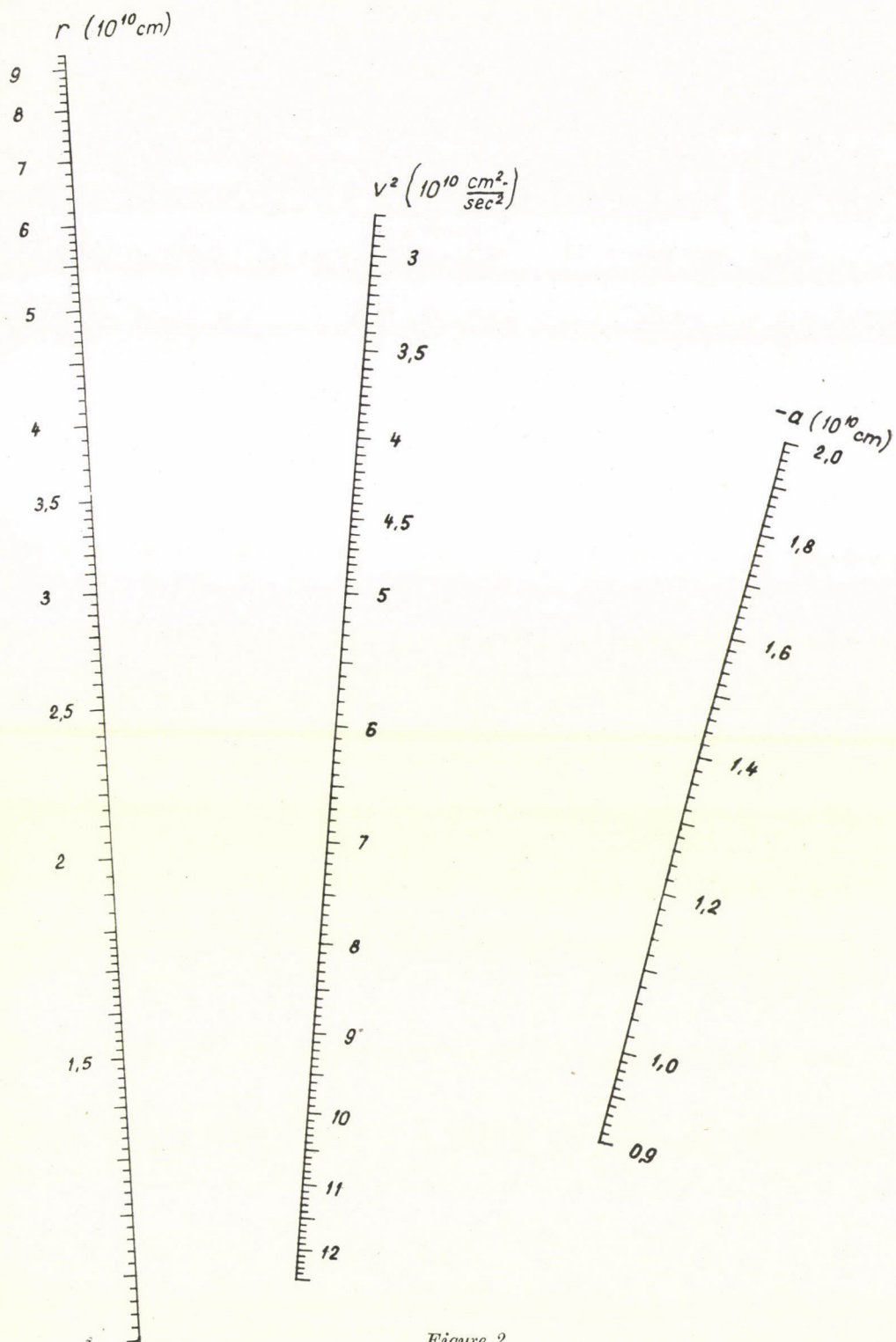
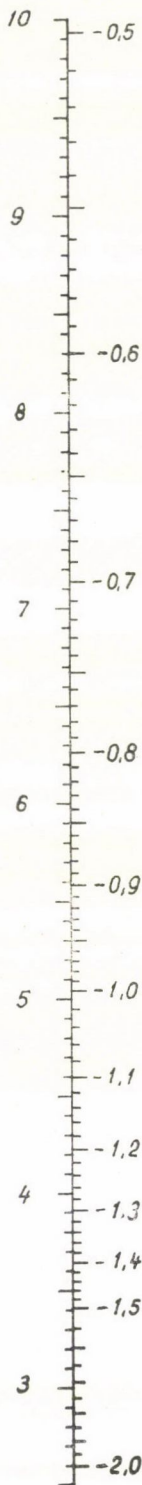


Figure 2.

$$\left(10^{10} \frac{\text{cm}^2}{\text{sec}^2}\right) V^2 \quad a(10^8 \text{cm})$$



where, in order to obtain the point P' with the ordinate $d' = a \cdot \sin u$, point P is projected not on the auxiliary circle, but on the hyperbola defined by the equation $x^2 - y^2 = a^2$. (The construction and the selection of the coordinate

system are shown in Fig. 5. The value of $\frac{d'}{a}$

may be computed from the obvious relationship

$$\frac{d'}{a} = \sqrt{\frac{x^2}{a^2} - 1} \text{ also numerically, where } x \text{ is the}$$

abscisse of point P .) In accordance, the eccentric

anomaly is obtained as $\text{ar sh } \frac{d'}{a}$. Nomograms 6a

and 6b have been constructed in order to dispense with hyperbolic function calculus. (Fig. 6a is applied to geocentric, Fig. 6b to selenocentric motion. They are used according to the key enclosed.) In practice starting from a point of known time co-ordinate first t_0 is determined from (5) then by applying the same equation again, the value t connected to any chosen point P may be obtained.

§ 6. Modification of the Velocity Vector when Passing Sphere Boundary

When the boundary of the Moon's sphere has been passed the selenocentric path may be constructed if the entrance velocity relatively to the Moon is known. According to [1] the latter may be found by vectorial subtraction of the actual velocity of the Moon from the geocentric velocity at the instant of the entrance of the rocket. When leaving the Moon's sphere the corresponding (momentary) velocity of the Moon shall be added to the selenocentric velocity, equal in absolute value to but different in direction from the entrance velocity, in order to obtain the geocentric velocity again. (Of course, at leaving the sphere of the Earth, the velocity of the rocket shall be increased by the momentary velocity of the Earth.)

Figure 3.

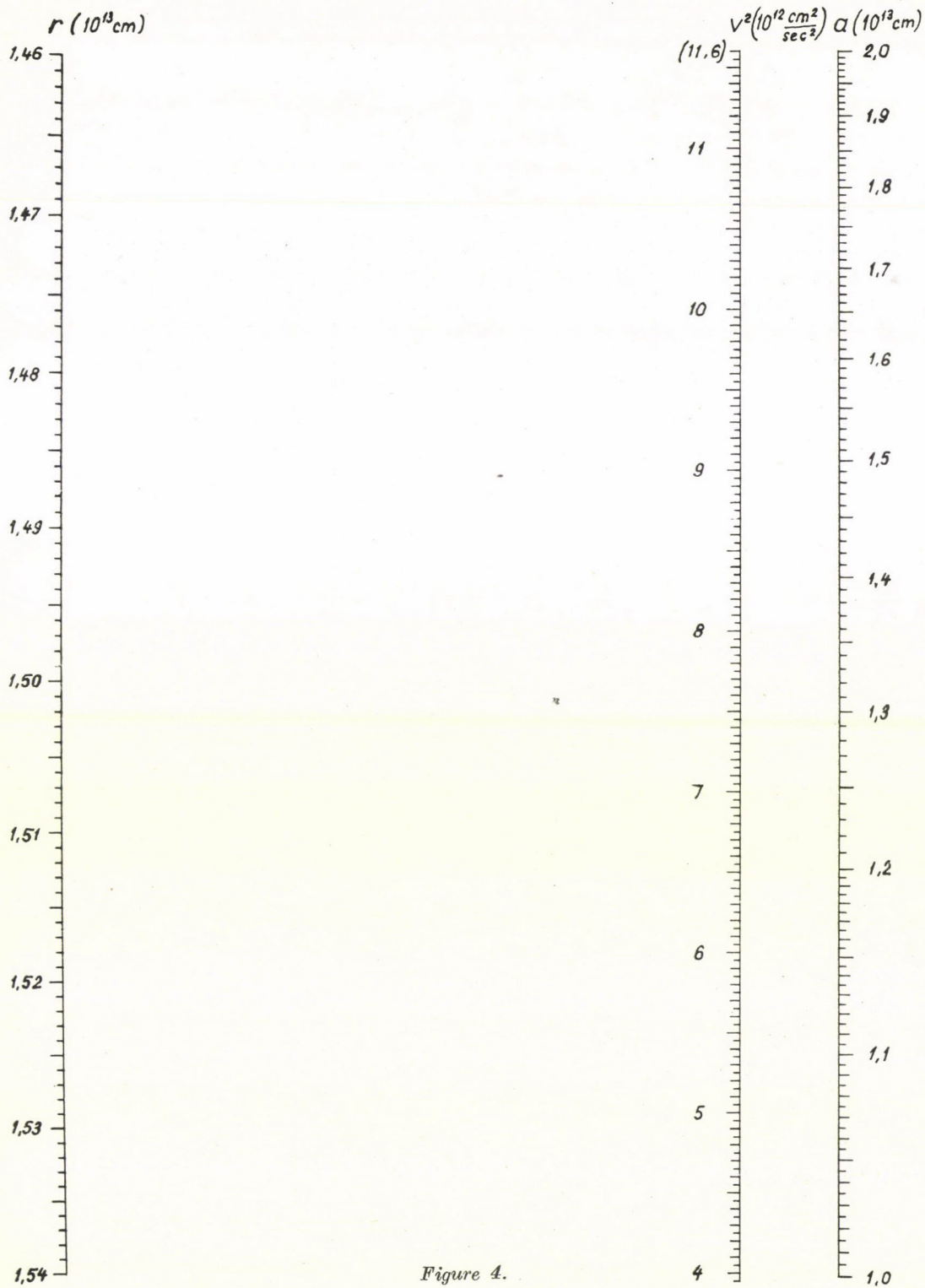


Figure 4.

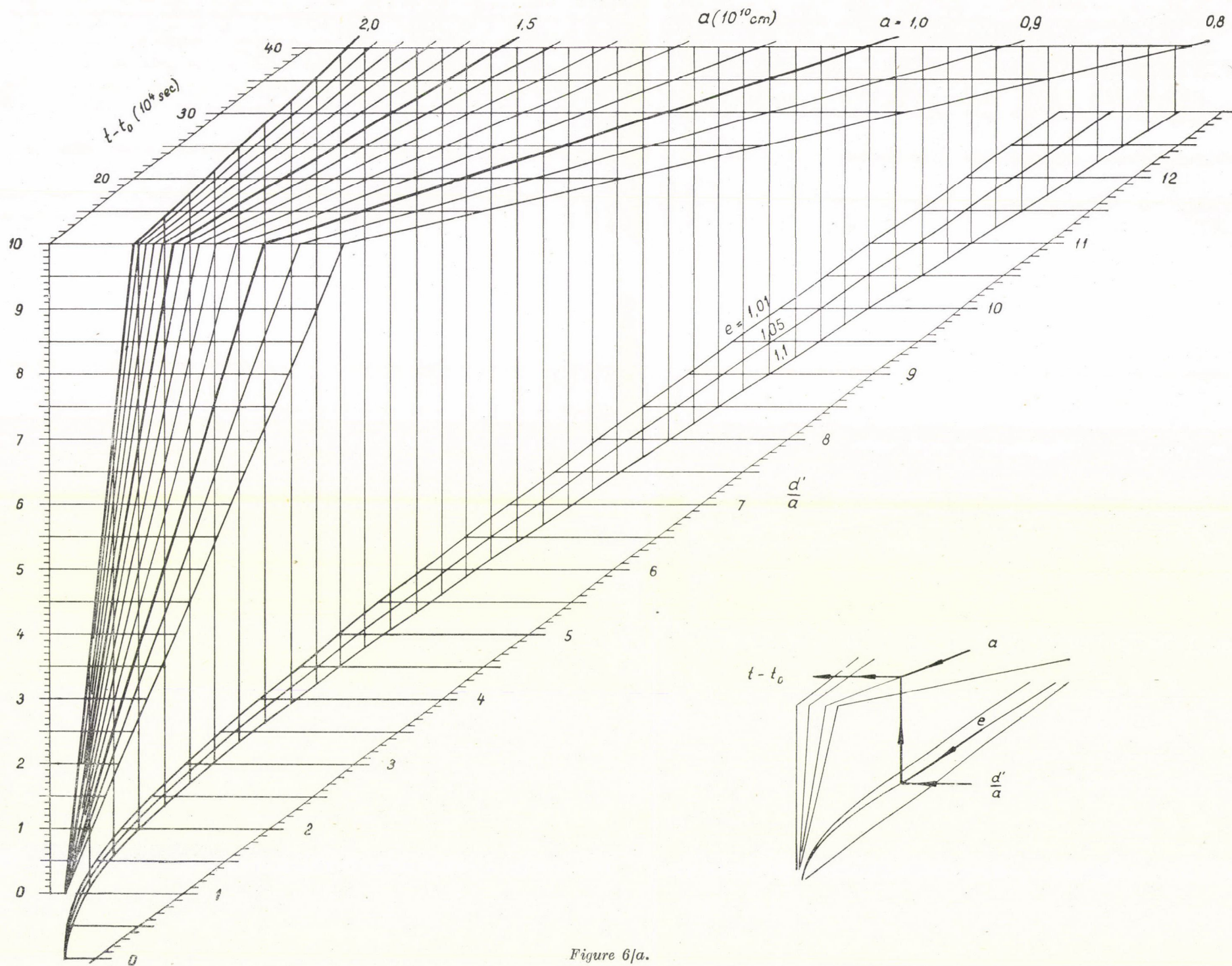


Figure 6/a.

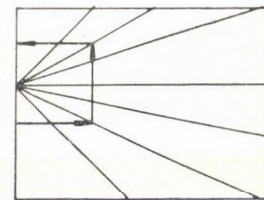
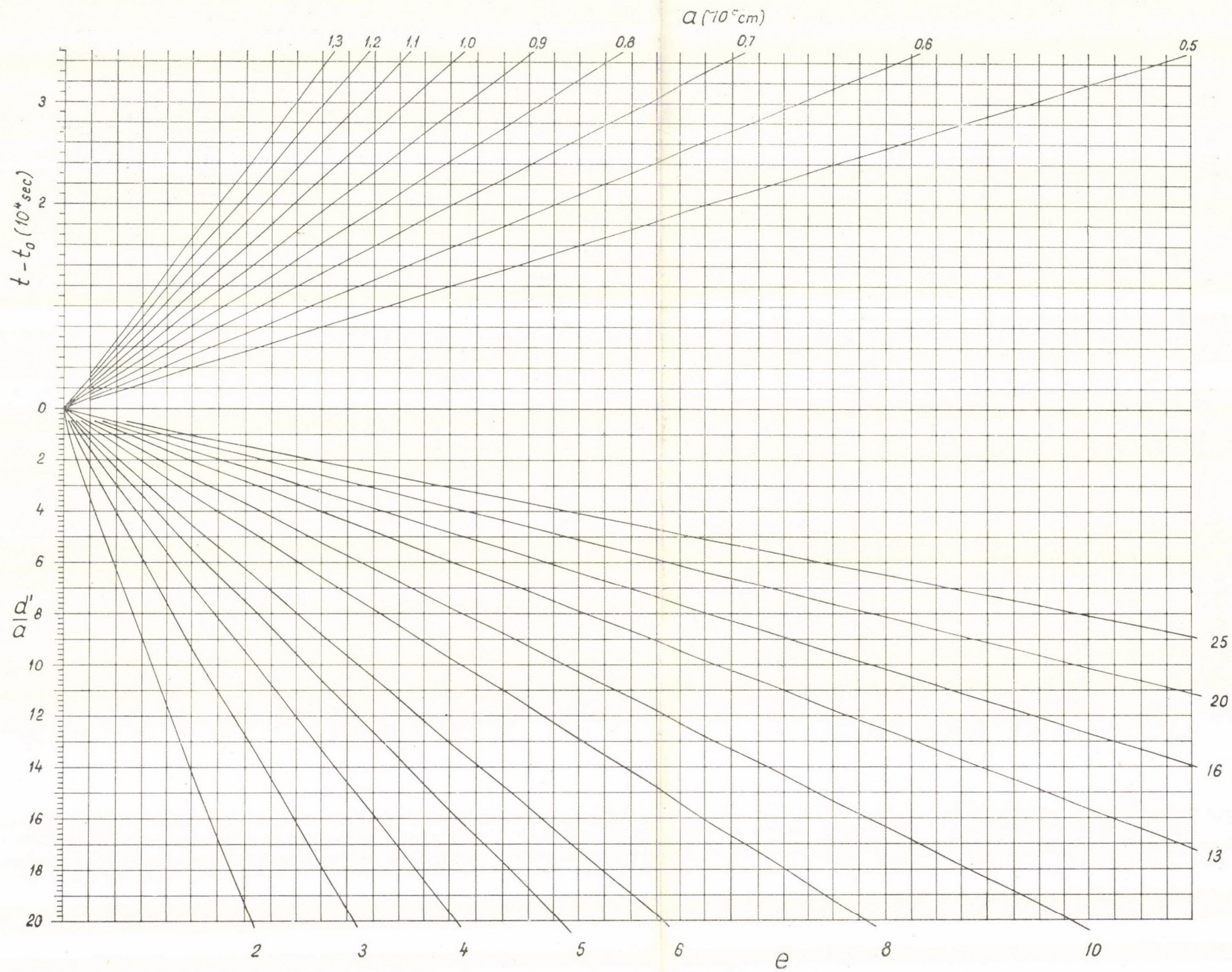


Figure 6/b.

where v_M denotes the velocity of the Moon, and v_r that of the rocket. According to our assumptions

$$\varrho^2 = n^2 + d^2$$

hence, substituting $\varrho = 66\,000$ km, τ may readily be found.

3. By subtracting the bigger one of the two roots from the moment of the fictitious intersection, the time wanted of entering the sphere is obtained. The position of the Moon at the same moment also is known, thus the intersection of the first hyperbola with the sphere of the Moon may be obtained.

Compute geocentric velocity at the boundary of the Moon's sphere from (3). Pass over to selenocentric velocity (II. §. 6) and starting from this, find the parameters of the second hyperbola (according to II. §§. 3 and 4). The construction also gives the nearest point to the Moon. The distance of the latter from the surface of the Moon and the time of its occurrence may easily be computed when the time of the entrance is known.

The point of exit is obtained by reflecting the point of entrance on the major axis of the hyperbola. As a consequence of symmetry, the absolute values of the selenocentric velocity at entrance and exit are equal. When the duration of the travel inside the Moon's sphere has been determined with the method described in II. §. 5, find the position of the Moon at the new instant and shift over the sphere together with the second hyperbola there parallelly. While discussing the present problem the system of co-ordinates affixed to the Moon may be considered as an inertia system.

After the exit, with the geocentric velocity computed according to II. § 6 construct the hyperbolic track as far as the boundary of the Earth sphere (II. §§. 3 and 4) and find the time of arrival (II. §. 5). Indicate the direction of the Sun at this instant and assume that the motion of the Earth is just perpendicular to this direction. By using $|v|$ as given by (3) determine the heliocentric velocity according to II. §. 6. Interpreting the direction of this velocity as an outer bisector of the angle, by adequate application of II. §§. 3 and 4 the parameters (a, e, c) of the elliptical path of the artificial planet may be determined. By the application of KEPLER's third law the period of rotation, and from (4) the instant of the perihelium may be determined.

IV. The Orbit of the Soviet Space Rocket Launched on 2nd January 1959

In order to check the accuracy of the method just described let us consider as an example the case of the first Soviet lunar probe.

Table I (Fundamental Data)

$h_1 = 2,09 \cdot 10^8$ km	$\lambda_1 = 40^\circ\text{W}$	$t_1 = 3. \text{ I. } 10^h$ U. T.
$h_2 = 2,37 \cdot 10^8$ km	$\lambda_2 = 86^\circ\text{W}$	$t_2 = 13^h$
$h_3 = 2,65 \cdot 10^8$ km	$\lambda_3 = 131^\circ\text{W}$	$t_3 = 16^h$
$h_4 = 2,84 \cdot 10^8$ km	$\lambda_4 = 160^\circ\text{W}$	$t_4 = 18^h$

Table II

	Notes	Comparative Data
At the start $t_s = 2. \text{ I. } 15^h 52^m$ $v_s = 10,2 \text{ km/s}$ $r_s = 7,9 \cdot 10^3 \text{ km}$ $\varphi_s = +21^\circ$ $\lambda_s = 140^\circ \text{ E}$	1 2 2	
First hyperbola $t_1 - t_s = 28^h 09^m$ $a_{01} = 1,05 \cdot 10^5 \text{ km}$ $e_{01} = 1,052$ $\alpha_{01} = 239^\circ$	3	
Entering Moon Sphere $t_1 = 3. \text{ I. } 20^h 01^m$ $v_1 = 2,53 \text{ km/s}$ $r_1 = 3,07 \cdot 10^5 \text{ km}$ $\varrho_1 = 6,6 \cdot 10^4 \text{ km}$	4 5	
Nearest point to Moon $t_M = 4. \text{ I. } 2^h 14^m$ $\varrho_M = 5,05 \cdot 10^3 \text{ km}$ $v_M = 2,86 \text{ km/s}$ $t_M - t_s = 34^h 22^m$	6	4. I. $2^h 59^m$ 5—6 10^3 km 34 ^h
Second hyperbola $t_2 - t_1 = 12^h 26^m$ $a_{12} = 7,42 \cdot 10^2 \text{ km}$ $e_{12} = 10,0$ $\alpha_{12} = 280^\circ$	3	
Exit from Moon's sphere $t_2 = 4. \text{ I. } 8^h 27^m$ $v_2 = 2,60 \text{ km/s}$ $r_2 = 4,41 \cdot 10^5 \text{ km}$ $\varrho_2 = 6,6 \cdot 10^4 \text{ km}$	5 4	
Third hyperbola $t_3 - t_2 = 67^h 47^m$ $a_{2,3} = 1,31 \cdot 10^5 \text{ km}$ $e_{2,3} = 1,084$ $\alpha_{2,3} = 186^\circ$	3	
Exit from Earth's sphere $t_3 = 7. \text{ I. } 4^h 14^m$ $v_3 = 1,97 \text{ km/s}$ $r_3 = 32,3 \text{ km/s}$ $d_3 = 9,30 \cdot 10^5 \text{ km}$ $\gamma_3 = 1,467 \cdot 10^8 \text{ km}$ $\gamma_3 = +1^\circ$	4 7 8	7/8. I. 32 km/s
Parameters of Art. $P = 447^d$ Planet's Orbit $a = 1,714 \cdot 10^8 \text{ km} = 1,145 \text{ A. U.}$ $e = 0,147$ $h_{ap} = 1,966 \cdot 10^8 \text{ km}$ $h_{per} = 1,462 \cdot 10^8 \text{ km}$ $\alpha = 17^\circ$ $t_{per} = 19. \text{ I.}$	9 10	450 ^d $1,715 \cdot 10^8 \text{ km}$ 0,148 $1,97 \cdot 10^8 \text{ km}$ $1,46 \cdot 10^8 \text{ km}$ 15° 14. I.

Notes to Table II

 t = time of transit (all times given in U. T.) r = distance from the Earth ϱ = distance from the Moon d = distance from the Sun v = velocity of the rocket

From the information available only the data contained in Table I have been used, and in accordance with this, for the construction of the starting hyperbola the procedure described in II. §. 3 has been employed. It is known, furthermore, that the engines of the rocket ceased to work at the height of 1500 km, therefore this point has been considered as the starting point of the orbit.

Our results have been listed in Table II. Comparative data have been taken from articles published in the PRAVDA up to the 12th January 1959 and from [5].

1. Starting height is 1500 km
2. Projection of the starting point on the surface of the Earth
3. The angle made by the major axis directed to the pericenter with the direction of the Vernal Equinox
4. Relatively to the Earth
5. Relatively to the Moon
6. Relatively to the surface of the Moon
7. Relatively to the Sun
8. The angle made by the direction of the heliocentric entrance velocity with the velocity vector of the Earth (the first one inclines towards the Sun)
9. The angle made by the major axis with the major axis of the orbit of the Earth
10. Perihelium transit time

The insignificant difference between the computed and communicated parameters of the track of the artificial planet 1 shows that the final results are influenced but insignificantly by the inaccuracies of the starting data and the construction. The rest of the known positions also lie near the track constructed. The soundness of the initial assumptions is supported by the interesting fact that the Moon was just in the ecliptic at the time of approach. (The inclination of the orbit of the artificial planet 1 has also been indicated as smaller than 1° .) The approximate orbit being known the approximate equatorial co-ordinates and distances necessary for the observation may readily be computed for any chosen instant.

The authors are indebted to Mr. A. BÉKÉSSY for his kind contribution in the construction of the nomograms.

(Received February 20, 1959.)

REFERENCES

- [1] ЕГОРОВ, В. А.: „О некоторых задачах динамики полета к Луне”. *Успехи Физических Наук* **63**.. (1957) 73—118.
- [2] СУББОТИН, М. Ф.: *Курс Небесной Механики*, Т. 2. (ОНТИ 1937) p. 194.
- [3] SMART, W. M.: *Celestial Mechanics* (London, 1953) p. 21.
- [4] CHARLIER, C. L.: *Die Mechanik des Himmels* (Leipzig, 1902) pp. 179—192.
- [5] *Nature*, No. 4654 (1959, I. 10), p. 83.

KÖZELÍTŐ MÓDSZER A HOLD MELLETT ELHALADÓ ŰRRAKÉTA PÁLYÁJÁNAK MEGSZERKESZTÉSÉRE

ALMÁR I. és BALÁZS B.

Kivonat

A dolgozat a „szféra-elv” ([1]) alkalmazásával a Hold mellett elhaladó rakéta mozgását szakaszokra bontja, amelyeken belül az egycentrumprobléma megoldásai alkalmazhatóak. Megadja a módszereket, melyek segítségével az űrrakéta egész pályája néhány, közvetlenül a fellövés után megadott pozíció alapján gyorsan szerkeszthető. A számolási lépések minimumra redukálása alkalmasan felvett nomogramok igénybevételel történhet (1, 2, 3, 4, 6. ábrák). Az eljárást az 1959. január 2.-i holdrakétára is alkalmazni lehetett; a kapott eredmények kielégítően egyeznek a később közzétett pályaadatokkal. (Kiindulási adatok: I. táblázat, eredmények: II. táblázat). A módszert megfigyelő állomások vezetői az űrrakéta pillanatnyi helyzetének közelítő meghatározására használhatják.

ПРИБЛИЖЕННЫЙ МЕТОД РАСЧЕТА ТРАЕКТОРИИ КОСМИЧЕСКОЙ РАКЕТЫ ПРОХОДЯЩЕЙ ВБЛИЗИ ЛУНЫ

I. ALMÁR и B. BALÁZS

Резюме

В данной работе применением «принципа сфер» ([1]) движение космической ракеты проходящей вблизи Луны разбивается на участки, а внутри отдельных участков применяются решения проблемы одного центра. Приводятся методы при помощи которых быстро может быть построена вся траектория космической ракеты на основании некоторых данных о положении сообщенных непосредственно после запуска ракеты. Снижение до минимума расчетных тактов осуществляется применением построенных соответствующим методом номограмм (рисунки № 1., 2., 3., 4., 6). Предлагаемый метод применялся и для расчета траектории космической ракеты запущенной 2-го января 1959 года; полученные результаты в достаточной мере совпадают с данными траектории позже сообщенными (исходные данные: таблица I, результаты: таблица II.). Предлагаемый метод может применяться руководителями наблюдательных станций для приближенного определения мгновенного положения космической ракеты.

**REMARK ON THE PAPER OF ALMÁR & BALÁZS :
„APPROXIMATE METHOD OF PLOTTING THE ORBIT OF
A SPACE ROCKET PASSING NEAR THE MOON”**

by

ANDRÁS BÉKÉSSY and KÁROLY TÓTH

In the course of the work of carrying out the nomograms of the above paper we noticed that the functions of the nomograms 6a and 6b can be relatively well represented even as alignment charts: moreover they can be handled easier and with the same (or with more) accuracy in this form. They are the nomograms 7 and 8.

On account of typographical reasons, we were compelled to use such small dimensions as they actually have. In order that they could be reproduced in larger dimension we give their equations :

Nomogram 7 :

$$\text{Scale for } t - t_0: \quad x = - \frac{960 (t - t_0)}{1 + 64 (t - t_0)^2} \quad y = \frac{180 (t - t_0)^2}{1 + 64 (t - t_0)^2}$$

$$\text{for } a: \quad x = \frac{1200 u}{100 + u^2}, \quad y = \frac{18000}{100 + u^2},$$

$$\text{where } u = \frac{1}{\sqrt{fm}} a^{3/2} \text{ with } fm = 3,98 \cdot 10^{20}$$

$$\text{for } \frac{d'}{a}: \quad x = \frac{12 \frac{d'}{a}}{1 + \frac{d'}{a} - 1,25 \operatorname{arsh} \frac{d'}{a}} \quad y = \frac{180 \left(1,09 \frac{d'}{a} - 1,25 \operatorname{arsh} \frac{d'}{a} \right)}{1 + \frac{d'}{a} - 1,25 \operatorname{arsh} \frac{d'}{a}}$$

$$\text{for } e: \quad x = - \frac{12}{1,25 e - 1}, \quad y = \frac{180 (1,25 e - 1,09)}{1,25 e - 1}$$

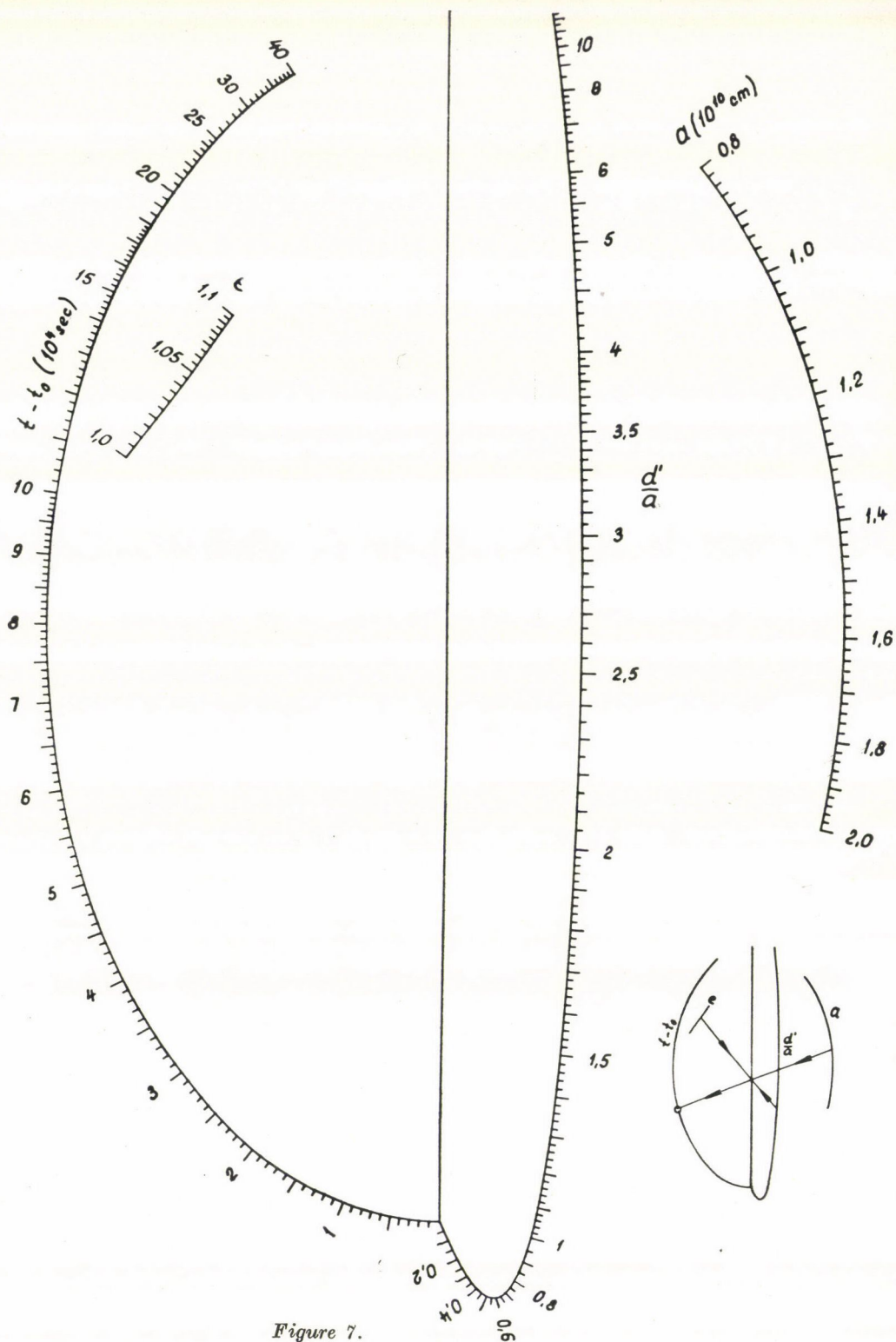


Figure 7.

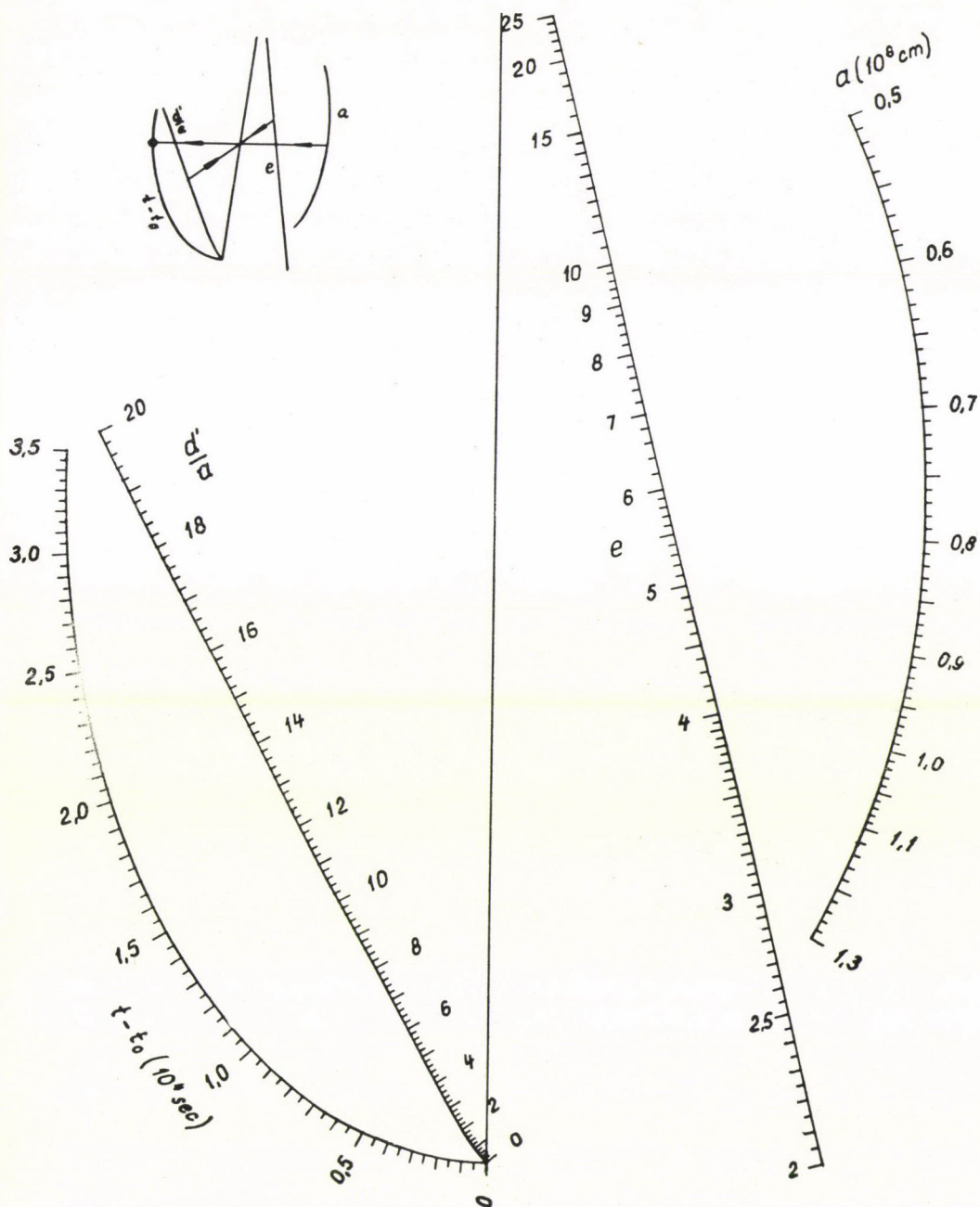
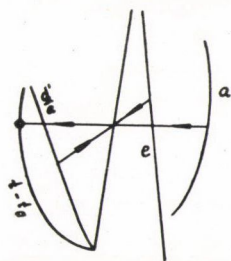


Figure 8.

Nomogram 8 :

$$\text{Scale for } t - t_0: \quad y = \frac{17,1 (t - t_0)^2}{1 + 0,09 (t - t_0)^2}$$

$$x = - \frac{36 (t - t_0)}{1 + 0,09 (t - t_0)^2}$$

$$\text{for } a: \quad x = \frac{3,6 u}{u^2 + 0,0009}$$

$$y = \frac{0,171}{u^2 + 0,0009}$$

$$\text{where } u = \frac{1}{\sqrt{fm}} a^{3/2} \quad \text{with } fm = 4,89 \cdot 10^{18}$$

$$\text{for } \frac{d'}{a}: \quad x = \frac{-1,68 \frac{d'}{a}}{1 - 0,018 \frac{d'}{a} - 0,009 \operatorname{arsh} \frac{d'}{a}}$$

$$y = \frac{3,42 \frac{d'}{a} - 1,71 \operatorname{arsh} v}{1 - 0,018 \frac{d'}{a} - 0,009 \operatorname{arsh} \frac{d'}{a}}$$

$$\text{for } e: \quad x = \frac{1,68 e}{0,009 e + 0,018}$$

$$y = \frac{1,71 e - 3,42}{0,009 e + 0,018}$$

**MEGJEGYZÉS ALMÁR I. ÉS BALÁZS B.: „KÖZELÍTŐ MÓDSZER A HOLD
MELLETT ELHALADÓ ŰRRAKÉTA PÁLYÁJÁNAK MEGSZERKESZTÉSÉRE”**

C. CIKKÉHEZ

BÉKÉSSY A. és TÓTH K.

Kivonat

A 7. és 8. ábra nomogramja pontsoros változatban ábrázolja ugyanazt a kapcsolatot, amely Almár I. és Balázs B. jelzett dolgozatában vonalsereges kivitelben szerepel (6a, 6b ábrák). Véleményünk szerint nomogramjaink a vonalsereges változatnál pontosabban olvashatók le.

**ЗАМЕЧАНИЕ К СТАТЬЕ I. ALMÁR И В. BALÁZS: „ПРИБЛИЖЕННЫЙ
МЕТОД РАСЧЕТА ТРАЕКТОРИИ КОСМИЧЕСКОЙ РАКЕТЫ
ПРОХОДЯЩЕЙ ВБЛИЗИ ЛУНЫ”**

A. BÉKÉSSY и К. TÓTH

Резюме

Номограммы 7. и 8. представляют с номограммами из выравненных точек те же номографические связи, которые в статье I. Almár и В. Balázs изображены со сетчатыми номограммами. По мнению авторов с настоящих номограммов величины снимаются точнее.

REMARKS ON A PAPER OF PÁL MEDGYESSY: "A MECHANICAL FUNCTIONAL SYNTHESIZER"

by

LÁSZLÓ CSORDÁS¹ and GYÖRGY MENCZEL¹)

I.

The mechanical functional synthesizer planned by P. MEDGYESSY²⁾ and operating in our Department has been used to calculate the projection of electron density in crystals.

Since the electron density in a crystal as well as its projection on a plane are periodical functions of the coordinates, they can be calculated by means of Fourier series [2]. By calculating the electron density in three dimensions one makes use of a three-dimensional series (with three sets of indices denoted by h , k and l while in projection the series is reduced to a two-dimensional one (one of the indices, say h being zero). Let us take for example the cristallographic axis x as the direction of projecting and project the unit cell on the yz plane. Supposed the crystal has a center of inversion, the projection of the electron density can be described in this case by the following Fourier series :

$$(1) \quad \varrho'(yz) = \frac{1}{A} \sum_k \sum_l F(0 \ kl) \cos 2\pi(ky + lz) = \frac{1}{A} \sum_k \sum_l F(0 \ kl) \cos (k\alpha + l\beta).$$

Notations :

A = the area of the projection

k, l = integers, indices of the series (corresponding to the "Miller indices" used in crystallography)

$F(0kl)$ = the coefficients of the series, which are identical with the "structure factors" of the crystallographic planes with Miller indices: $(0kl)$. This quantities can be measured on X-ray diffraction photographs.

y, z = relative coordinates referred to the length of the corresponding edge of the unit cell.

$\varrho'(yz)$ = electron density measured in units: number of electrons per square Ångström.

$$\alpha = 2\pi y$$

$$\beta = 2\pi z$$

¹⁾ Department of Physics No. I., Eötvös Loránd University.

²⁾ Pál Medgyessy: "A Mechanical Functional Synthesizer" (*Publications of the Mathematical Institute of the Hungarian Academy of Science*. 2 (1957) 33-42.)

This two-dimensional synthesis can be reduced to two successive one-dimensional ones by means of a simple trigonometric substitution (For sake of simplicity we introduce the notation: $\varrho(yz) = A\varrho'(yz)$)

$$(2) \quad \varrho(yz) = \sum_k \left[\sum_l F(0kl) \cos l\beta \right] \cos k\alpha - \sum_k \left[\sum_l F(0kl) \sin l\beta \right] \sin k\alpha.$$

Introducing for the results of the first summation the notation:

$$S_1(k, z) = \sum_l F(0kl) \cos l\beta; \quad S_2(k, z) = \sum_l F(0kl) \sin l\beta$$

and writing the constant terms of the series: $F(000)$ and $S_0 = S_1(0, z)$ separately:

$$(3) \quad \varrho(yz) = F(000) + S_0 + \sum_{k \neq 0} [S_1(k, z) \cos k\alpha - S_2(k, z) \sin k\alpha].$$

Since we want to work with the mechanical synthesizer we have to keep in mind that it has patterns of the form

$1 \pm \sin kx$ resp. $1 \pm \cos kx$ and it can synthesize terms like:

const. $(1 \pm \sin kx) \cdot$ resp. const. $(1 \pm \sin kx)$

where the constant factor is always positive. So it is comfortable for us to transform our equation:

$$(4) \quad \varrho(yz) = \sum_{k \neq 0} [|S_1(k, z)| + S_1(k, z) \cos k\alpha + |S_2(k, z)| - S_2(k, z) \sin k\alpha] - \\ - \left\{ \sum_{k \neq 0} [|S_1(k, z)| + |S_2(k, z)|] - F(000) - S_0 \right\}$$

Introducing the notation:

$$\Sigma = \sum_{k \neq 0} [|S_1(k, z)| + S_1(k, z) \cos k\alpha + |S_2(k, z)| - S_2(k, z) \sin k\alpha]$$

$$C = \sum_{k \neq 0} [|S_1(k, z)| + |S_2(k, z)|] - F(000) - S_0$$

we obtain the equation

$$(5) \quad \varrho(yz) = \Sigma - C$$

from which it is obvious that the projection of the electron density $\varrho(yz)$ is obtained by subtracting the constant term C from the value Σ indicated by the mechanical synthesizer.

The result of a summation appears on the instrument as end-points of bars, in 37 equidistant points having the values $0 < \alpha < \pi$; in our case, this sum corresponds to the projection of the electron density along a line $z = \text{constant}$ and $0 < y < 1/2$. (see the paper of P. MEDGYESSY [1].)

At first time, after having read the positions of the end-points of bars on a scale fixed under the bars, we wrote the data on an yz net. After this

we put a transparent paper over the net and traced the lines of equal electron density in order to find the maxima, that is the places of atomic centers.

In order to speed up our work we have initiated a new method for fixing the results of the second summation. Instead of reading, we draw the sum on a leaf of drawing-paper: First, we trace a straight line on the paper then we put the paper between the bars and the scale so that the straight line coincided with the value

$$C = \sum_{\substack{k \\ k \neq 0}} [|S_1(k, z)| + |S_2(k, z)|] - F(000) - S_0,$$

then we mark the end of each bar and cut the figure we have got in this way. Such a figure corresponds to the projection of the electron density along a line $z = \text{const.}$ and $0 < y < \frac{1}{2}$. As we have calculated the first sums $S_1(k, z)$ and $S_2(k, z)$ in equidistant points of the z axis, we shall have now as many figures as z values.

We constructed a square frame for the figures and mounted them into the inner slits of the frame. The figures were arranged according to their z values, while the O lines of each figure were held in equal height. So we obtained a very clear three dimensional representation of the function $\varrho(yz)$ (Fig. 1.). Two horizontal directions correspond here to the y resp. z axis and the vertical direction to $\varrho(yz)$. The maxima are immediately observable and their coordinates and height could be read on a scale made of plexiglass.

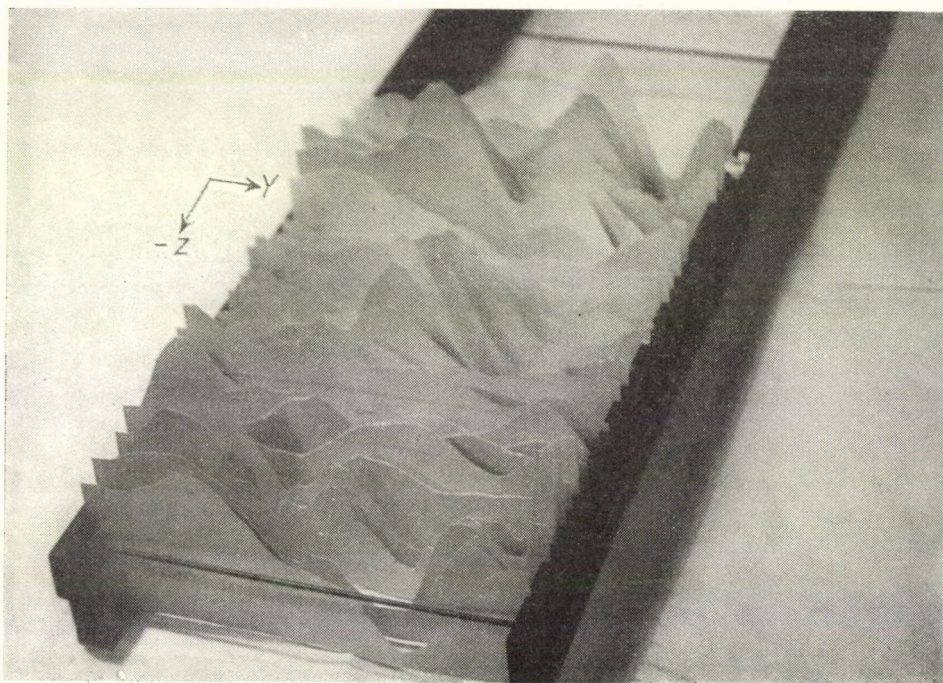


Figure 1.

The main advantage of the new method beside its graphic character is, that it reduces the working time. With the old method the calculations took up scarcely less time than with the BEEVERS—LIPSON strips. Now we work three times faster and we do not need drawing an electron density map.

II.

As mentioned in section I. the instrument has patterns of the shape $1 \pm \sin k$ and $1 \pm \cos k$. The values on the patterns are calculated in steps of 5 degrees and also the sum is obtained in these points. In the course of crystal structure work, in case of a large unit cell, there occur indices of 30—40 and even larger ones. It is easy to show that the preparing of the patterns is needed only up to $k = 18$ and patterns with $k > 18$ can be substituted by patterns with $k < 18$, *supposed we want to obtain the results in the mentioned equidistant points of 5 degrees.*

It is suitable to introduce the following notation: We denote "pattern $\sin k_i$ " and "pattern $\cos k_i$ " the patterns of the shape $1 + \sin k_i$ and $1 + \cos k_i$ further "pattern $-\sin k_i$ " and "pattern $-\cos k_i$ " the patterns of the shape $1 - \sin k_i$ resp. $1 - \cos k_i$. As "even steps" are denoted the steps $\alpha = 2n5^\circ$ of the patterns and as "odd steps" those of $\alpha = (2n + 1)5^\circ$. We shall use the name "even bars" and "odd bars" for the bars shifted by the even resp. odd steps of the patterns.

Keeping in view that we calculate in steps of 5 degrees and that the values of the trigonometric functions of arguments larger than 90 degrees can be substituted by values belonging to arguments less than 90 degrees, we can fit together following relations:

- a) $\sin [(2n + 1) (36 - k) 5^\circ] = \sin [(2n + 1) k 5^\circ]$
- b) $\sin [2n (36 - k) 5^\circ] = -\sin (2n k 5^\circ)$
- c) $\cos [(2n + 1) (36 - k) 5^\circ] = -\cos [(2n + 1) k 5^\circ]$
- d) $\cos [2n (36 - k) 5^\circ] = \cos (2n k 5^\circ).$

Making use of the previous and similar relations the following Table I. can be prepared. This table shows the sign relationship between $k > 18$ and the substituting patterns $k < 18$.

Table I

α		k	I $36 - k$	II $k - 36$	III $72 - k$
Sin	$(2n+1) \cdot 5^\circ$	+	+	—	—
	$2n \cdot 5^\circ$	+	—	+	—
Cos	$(2n+1) \cdot 5^\circ$	+	—	—	+
	$2n \cdot 5^\circ$	+	+	+	+

The column I. is to be used if $18 \leq k \leq 36$; column II. if $36 \leq k \leq 54$ and column III. if $54 \leq k \leq 72$. (The case $k > 72$, has no practical significance.)

We can elucidate the use of the Table on an example: It is to be added $S_2(21, z) \cdot \sin 21 \alpha$. From the table it is clear that instead of the pattern sin 21 we must apply pattern sin 15 for the odd steps, and pattern $-\sin 15$ for the even steps. In practice this can be achieved by putting pattern sin 15 into the instrument and shift forward so that only the odd bars moved and the even ones remained in their original positions. After this we put in pattern $-\sin 15$ and now we must be able to set in motion only the even bars and leave the others unmoved.

Our instrument is built so that the bars can move only together. We must transform the instrument so that the even bars and the odd bars could move either separately or together. By a slight modification this can be done but the description of the technical details exceed the scope of this communication.

We have to treat particularly the function: $\cos 36 \alpha$. It has values $+1$ at $\alpha = 2n \cdot 5^\circ$ and -1 at $\alpha = (2n + 1) \cdot 5^\circ$. According to our notation in section I. its coefficient is $S_1(36, z)$ and the value of the function $S_1(36, z) \cos 36$ is at the even steps equal to $S_1(36, z)$ and at the odd steps: $-S_1(36, z)$. According to Table I. we need the pattern $\cos 0$ for its addition to the series. This pattern has been eliminated by including his coefficient S_0 into the constant term C (equation 5. in section I.), therefore it would be necessary to prepare pattern $\cos 36$ as the unique one from among the patterns $k > 18$. Still there is a simple way to avoid this, by including also this term into the constant term of the addition. In practice this can be performed in the following way: After having added all the terms included in Σ (equation [5], section I.) we put our leaf of paper under the bars so that the $\varrho(yz) = 0$ line coincided with the value $C' = C - S_1(36, z)$ of the scale and mark the end-points of the even bars. After this, we shift the O line to the value $C'' = C + S_1(36, z)$ and now we mark the positions of the odd bars.

It has to be mentioned that the Beevers—Lipson strips applied in crystal structure analysis are constructed on the basis of similar ideas. Here, the values of the trigonometric functions are printed on both sides of a strip of paper, in intervals of 3 degrees. The values belonging to the even multiples of 3 degrees are on one side and those belonging to the odd multiples on the other side.

(Received December 22, 1958.)

REFERENCES

- [1] MEDGYESSY P.: "A Mechanical Functional Synthesizer". *Publ. of the Mathematical Institute of the Hungarian Academy of Science*. 2 (1957) 33—42.
- [2] LIPSON, H.—COCHRAN, W.: *The Determination of Crystal Structures*. The Crystalline State Vol. III. London 1953.

MEGJEGYZÉSEK MEDGYESSY PÁL: „EGY MECHANIKAI FÜGGVÉNYSZINTETIZÁTOR” CÍMŰ CIKKÉHEZ

CSORDÁS L. és MENCZEL Gy.

Kivonat

Medgyessy Pál [1] cikkében leírt szintetizátorral számoltunk. A számolási időt kb. harmadára csökkentettük azáltal, hogy az eredményeket nem olvastuk le, hanem a pálcák alá karton papírost téve megjelöltük a pálcák végpontjait és az így kapott ábrát kivágtuk. Az ábrákat egymás mellé rakva az elektronsűrűség vetületi eloszlásának szemléletes, háromdimenziós képe alakult ki.

Kimutatjuk, hogy a sablonok $h = 18$ -ig való elkészítésével az összegezési lehetőséget bármilyen magas frekvenciáig ki lehet terjeszteni. Ez által a sablonok költséges elkészítését megtakaríthatjuk.

ЗАМЕЧАНИЯ К СТАТЬЕ Р. MEDGYESSY: „ОДИН МЕХАНИЧЕСКИЙ ФУНКЦИОНАЛЬНЫЙ СИНТЕТИЗАТОР”

L. CSORDÁS и Gy. MENCZEL

Резюме

Мы считали с помощью синтетизатора, описанного в статье Р. Мед-
gyessy [1]. Мы сократили время счёта примерно в три раза следующим
образом: мы не читали результаты, а, положив под палочки картон, отме-
тили концы палочек и вырезали полученный чертёж. Поместив рядом по-
лученные чертежи, мы получили наглядную, трёхмерную картину проек-
ционного распределения плотности электронов.

Мы показываем, что изготовляя шаблоны до $h = 18$ возможность
суммирования может быть распространена до как угодно высокой частоты.
Таким образом мы можем сэкономить дорогостоящее изготовление шаб-
лонов.

AUTOKLÁVOK SOROS ÉS PÁRHUZAMOS KAPCSOLÁSÁRÓL ÉS A KEVERÉS ELMÉLETÉRŐL

RÉNYI ALFRÉD

E dolgozat első négy paragrafusában az Almásfüzitői Timföldgyár által felvetett probléma megoldását ismertetjük. Az 5. §. néhány megjegyzést tartalmaz a keverés általános elméletére vonatkozólag.

1. §. A probléma megfogalmazása

Adva van r számú autokláv, amelyek egymás után vannak kapcsolva. Az első autoklávhoz óránként v köbméter anyag (lúgos oldatban elkevert bauxitpor) áramlik be és ugyanennyi anyag áramlik tovább óránként az első autoklávból a másodikba, a másodikból a harmadikba, s. í. t., és végül ugyanennyi áramlik ki az utolsó, vagyis az r -edik autoklávból. Az autoklávok teljesen egyformák; mindegyik V köbtartalmú. Az autoklávokban az anyagot állandóan keverik. Első közelítésben feltesszük, hogy a keverés tökéletes, vagyis az egyes autoklávokba beáramló anyag gyakorlatilag azonnal teljesen elkeverődik.

Meghatározandó, hogy az autoklávokon átáramló anyag hányadrésze tölt a rendszerben t óránál rövidebb időt. Ezt elsősorban a $t = t_0$ esetre kell meghatározni, ahol t_0 azt az időtartamot jelöli, amennyi szükséges ahhoz, hogy a kívánt kémiai reakció végbemenjen. Jelölje $F_r(t)$ azt, hogy az anyag hányadrésze tölt a rendszerben t óránál rövidebb időt. A feladat tehát az $F_r(t)$ függvény kiszámítása.

2. §. A probléma valószínűségszámítási interpretációja

Az átáramló anyag egy-egy részecskéjének¹ átfutási ideje a véletlentől függ, attól ugyanis, hogy a keverés során mi történik az illető részecskével. Ez a mennyiség tehát valószínűségi változónak tekinthető. Jelöljük ezt a valószínűségi változót ξ_r -rel. Az egész rendszeren való átfutás időtartama az egyes autoklávokon való átfutási idők összegével egyenlő. Ha tehát ξ_j jelöli a j -edik autoklávon való átfutás időtartamát, akkor $\zeta_r = \sum_{j=1}^r \xi_j$. Mivel az egyes autoklávok egyformák, a $\xi_1, \xi_2, \dots, \xi_r$ valószínűségi változók egyforma eloszlásúak. Nyilvánvaló továbbá, hogy a ξ_j ($j = 1, 2, \dots, r$) valószínűségi változók függetlenek egymástól, mivel az egyes autoklávokban

¹ A probléma bizonyos leegyszerűsítését jelenti az, hogy egyes részecskékről beszélünk, hiszen valójában a bauxit-por szemcséi szétaprózódnak az autoklávban. Ez a körülmény azonban nincs lényeges befolyással az átfutási időre, hiszen a bekerülő nagyobb szemcséket úgy tekinthetjük, mintha azok nagyszámú kis részecskékből állnának, amelyek az autoklávban elválhatnak egymástól.

lejátszódó keverési folyamatok egymástól függetlenül mennek végbe. Ki fogjuk mutatni, hogy a tökéletes keverés feltételezése mellett a ξ_j valószínűségi változók exponenciális eloszlásúak, vagyis, ha $F_1(t)$ jelöli annak a valószínűségét, hogy egy molekula egy autoklávban t -nél rövidebb időt tölt (minden időtartamot órában fejezünk ki), akkor

$$(1) \quad F_1(t) = 1 - e^{-\lambda t},$$

ahol $\lambda = \frac{v}{V}$. (V köbméterekben és v m³/órákban értendő, $\frac{1}{\lambda}$ dimenziója óra.)

Ez a valószínűségi számítás bizonyos általános tételeiből következik, de belátható teljesen elemileg is. Az alábbi szemléletes megfontolás azok számára szól, akik a valószínűségi számításban kevésbé járatosak és minden előismeret nélkül kívánnak az (1) képlet helyességéről meggyőződni.

Tegyük fel, hogy az anyag minden m³-ében N számú részecske van; akkor az autoklávban egyidejűleg NV számú részecske van. 1 óra alatt Nv számú részecske hagyja el az autoklávot, tehát t óra alatt Nvt számú részecske távozik az autoklávból. A tökéletes keverés feltevéséből következik, hogy egy adott pillanatban az autoklávban tartózkodó összes részecske ugyanolyan valószínűséggel távozhat *elsőnek*. Ez a valószínűség tehát $\frac{1}{NV}$. Annak

a valószínűsége tehát, hogy a vizsgált részecske ne elsőként távozzék, $1 - \frac{1}{NV}$.

Kérdés, mi a valószínűsége, hogy egy kiválasztott részecske pontosan k -adikként távozzon. Ez a valószínűség nyilván $\left(1 - \frac{1}{NV}\right)^{k-1} \frac{1}{NV}$, hiszen az, hogy a szóbanforgó részecske k -adikként távozik, azt jelenti, hogy sem elsőnek, sem másodiknak, ..., sem $(k-1)$ -ediknek nem távozik, a következő lépésben azonban távozik. Ilyenmódon annak a valószínűsége, hogy egy részecske egy autoklávban t -nél rövidebb időt töltsön,

$$(2) \quad F_1(t) = P(\xi_k < t) = \sum_{k=1}^{Nvt} \left(1 - \frac{1}{NV}\right)^{k-1} \frac{1}{NV} = 1 - \left(1 - \frac{1}{NV}\right)^{Nvt}.$$

Mivel NV igen nagy, $1 - \frac{1}{NV}$ helyett jó közelítéssel $e^{-\frac{1}{NV}}$ írható. Ennél fogva

$$(3) \quad F_1(t) \sim 1 - e^{-\frac{vt}{V}} = 1 - e^{-\lambda t}, \quad \text{ahol } \lambda = \frac{v}{V}$$

amit bizonyítani akartunk.²⁾

²⁾ Egy másik, szintén igen egyszerű bizonyítás a következő: a tökéletes keverés definíciójából következik, hogy annak a valószínűsége, hogy egy részecske, amely az s időpontban az autoklávban van, az $s + t$ időpont előtt távozzék, csak t -től függ és nem függ attól, hogy az illető részecske az s időpontot megelőzőleg milyen hosszú időt töltött az autoklávban. Ebből rögtön következik, hogy ha $G(t) = 1 - F_1(t)$, akkor $G(t)$ monoton csökkenő és eleget tesz a $G(s + t) = G(s)G(t)$ függvényegyenletnek. E függvényegyenletnek azonban, mint jól ismeretes, az egyedüli monoton csökkenő megoldásai a $G(t) = e^{-\lambda t}$ függvények, ahol $\lambda > 0$.

3. §. A probléma megoldása

Határozzuk meg most ζ_r eloszlását. Egy jólismert tétel szerint (lásd pl. [1], 230—233. o.) r számú független és egyforma exponenciális eloszlású valószínűségi változó összegének eloszlása r -edrendű gamma-eloszlású, tehát $F_r(t)$ -vel jelölve ζ_r eloszlásfüggvényét (vagyis annak a valószínűségét, hogy egy részecske t óránál rövidebb idő alatt fusson át az r autokláv-ból álló rendszeren) fennáll a következő összefüggés:

$$(4) \quad F_r(t) = \frac{1}{(r-1)!} \int_0^{\lambda t} x^{r-1} e^{-x} dx.$$

$F_r(t)$ tehát ún. nem-teljes gamma-függvény, mégpedig bevezetve a

$$(5) \quad \Gamma_r(y) = \frac{1}{\Gamma(r)} \int_0^y u^{r-1} e^{-u} du$$

jelölést, ahol $\Gamma(r)$ a jólismert gamma-függvény,

$$(6) \quad \Gamma(r) = (r-1)! = \int_0^{\infty} x^{r-1} e^{-x} dx,$$

akkor

$$(7) \quad F_r(t) = \Gamma_r(\lambda t), \quad \text{ahol } \lambda = \frac{v}{V}.$$

A nem-teljes gamma-függvény tabulálva van. (Egy rövid táblázata megtalálható pl. [1] 139—141. oldalain).

Lássunk egy számpéldát. Legyen $V = 32 \text{ m}^3$, $v = 40 \text{ m}^3/\text{óra}$, $r = 7$ és $t = 3 \text{ óra}$, akkor $\lambda = 1,25 \text{ óra}^{-1}$ és így a 7 autokláv-ból álló rendszerben az átáramló anyagnak csak $\Gamma_7(3,75) = 0,087$ része, azaz 8,7%-a tölt 3 óránál rövidebb időt.

4. §. Az autokláv-park leggazdaságosabb kihasználásának problémája

Ha óránként $A \text{ m}^3$ anyagot akarunk átáramoltatni és összesen R autokláv áll rendelkezésre, akkor kérdés, milyen csoportokra célszerű elosztani az autoklávokat.

A következő elrendezéseket vizsgáljuk: az autokláv-parkot egyenként r sorbakapcsolt autokláv-ból álló csoportra osztjuk és az így nyert $\frac{R}{r}$ csoportot

párhuzamosan kapcsoljuk; r -ről feltesszük, hogy R egy osztója. Kérdés, hogy a rendszer hatásfoka milyen elrendezés mellett lesz maximális? Másszóval ha t_0 időtartamig kell egy-egy részecskének az autokláv-rendszerben tartózkodnia ahhoz, hogy a kívánt reakciók végbemenjenek, r mely értékénél lesz minimális az átáramló anyagnak az a hányada, amely t_0 óránál kevesebbet töl a rendszerben? A kérdésre a válasz nem könnyű, hiszen ha több autoklávot kapcsolunk sorba, vagyis r értékét növeljük, akkor egy-egy csoportban azono

áramlási sebesség mellett hosszabb időt töltene az anyag, azonban ekkor az áramlás sebessége is arányosan megnő, hiszen egy autokláv-csoporton nyilván $\frac{Ar}{R}$ m³ anyag áramlik át. A rendszer hatásfokára tehát két ellentétes irányú tényező hat. Ilyen módon általában az várható, hogy r -nek sem a túl kicsiny, sem a túl nagy értékei nem kedvezőek. A rendszer hatásfokát a mondottak alapján kézenfekvő azzal mérni, hogy az anyag mekkora hányada tölt a rendszerben az előírt t_0 kritikus időtartamnál hosszabb időt. Tehát a szóbanforgó kapcsolat mellett a rendszer hatásfoka

$$(8) \quad H_r = 1 - \Gamma_r \left(\frac{Ar t_0}{RV} \right).$$

Ennélfogva r értékét minden adott esetben (adott A , t_0 , R és V értékek mellett) úgy kell megválasztani, hogy $1 - \Gamma_r \left(\frac{Ar t_0}{RV} \right)$ maximális legyen, azaz úgy, hogy $\Gamma_r \left(\frac{Ar t_0}{RV} \right)$ minimális legyen.

Mivel a számbajövő r értékek száma kicsiny, a nem-teljes Γ -függvény táblázata segítségével ez a számítás gyorsan és könnyen elvégezhető. Elvégezhető az optimális r -érték meghatározása grafikusan is; ehhez nyilván csak a $\Gamma_r(x)$ függvények grafikonjaira van szükség.

Vizsgáljuk most meg a szóbanforgó elrendezés mellett a ζ_r átfutási idő várható értékét és szórását. Mivel a $\Gamma_r(\lambda x)$ eloszlásfüggvényű gamma-eloszlás várható értéke, mint ismeretes, $\frac{r}{\lambda}$ és szórása $\frac{\sqrt{r}}{\lambda}$, tehát ζ_r várható értéke

$$(9) \quad \mathbf{M}(\zeta_r) = r \cdot \frac{RV}{Ar} = \frac{RV}{A}.$$

Más szóval az átfutási idő várható értéke független r értékétől. Ha például $R = 24$, $V = 32$ m³ és $A = 120$ m³/óra, akkor az átfutási idő várható értéke a számításba jövő $r = 1, 2, 3, 4, 6, 8, 12, 24$ értékek mindegyikére 6 óra 24 perc. Az átfutási idő szórása a mondottak szerint

$$(10) \quad \mathbf{D}(\zeta_r) = \sqrt{r} \frac{RV}{Ar} = \frac{RV}{A \sqrt{r}},$$

a szórás tehát r növekedésével csökken. Ha az adatok olyanok, hogy az átlagos átfutási idő, tehát $\frac{RV}{A}$ lényegesen nagyobb az előírt kritikus időtartamnál, akkor tehát célszerű r értékét minél nagyobbra választani.

A gyakorlatban az optimális kapcsolat kiválasztásánál még egy szempontra figyelemmel kell lenni. Az autoklávokat ugyanis rendszeresen karban kell tartani, és így az a célszerű, ha az autokláv-csoportok közül egy ciklikusan mindig üzemben kívül van, a szükséges karbantartási munkálatok elvégzése céljából. Az optimális elrendezés kiválasztása a közölt képletek alapján ezen gyakorlati szempont szem előtt tartásával is könnyen elvégez-

hető. Ha összesen R autokláv van egy üzemben és ezeket r autoklávból álló csoportokba osztjuk, úgy, hogy ezen csoportok közül felváltva egy mindig üzemben kívül van, akkor a működő csoportok száma $\frac{R}{r} - 1$, tehát $v = \frac{rA}{R-r}$ és így r értékét úgy kell kiválasztani R osztói közül, hogy $\Gamma_r\left(\frac{Art_0}{V(R-r)}\right)$ minimális legyen.

5. §. A nem tökéletes keverés problémája

A valóságban a keverés sohasem tökéletes a fenti értelemben, vagyis nem igaz az, hogy minden egyes részecske, amely egy adott időpontban az autoklávban tartózkodik, ugyanolyan valószínűséggel hagyja el elsőnek az autoklávot, hanem a keverés véges sebessége folytán a kevéssel azelőtt az autoklávba beáramlott részecskék kisebb valószínűséggel távoznak, mint a már régebben az autoklávba bekerült részecskék. Ennek következtében egyes részecskék által egy-egy autoklávban töltött időtartam nagyobb lesz, mint a tökéletes keverés mellett volna, míg más részecskék esetében kisebb lesz. (Az átlagos átfutási idő persze nem változik, hiszen az szükségképpen megegyezik azzal az időtartammal, amelyet a részecskék az autoklávban töltenének, ha mindegyik ugyanannyi időt töltene az autoklávban.)

Ha figyelembe kívánjuk venni, hogy a keverés nem tökéletes, akkor valamilyen módon meg kell határozni a keverés *intenzitását*. A szóbanforgó probléma szempontjából természetes mértékszám a keverés intenzitásának az egyes részecskék egy-egy autoklávon való átfutási idejének relatív szórásnégyzete tehát az

$$(11) \quad i = \frac{D^2(\xi)}{M^2(\xi)}$$

hányados. Ha a keverés tökéletes, akkor ξ exponenciális eloszlású és így $i = 1$. Ha keverés egyáltalán nincsen, akkor az átfutási idő állandó és így $D^2(\xi) = 0$, tehát $i = 0$. Kérdés, milyen lesz nem-tökéletes keverés mellett az átfutási idő eloszlása. Erre vonatkozólag plauzibilis feltevés, hogy nem tökéletes keverés esetében az *egy* autoklávon való átfutási idő gamma-eloszlású lesz.

Ez a hipotézis annak a fizikai képnek felel meg, hogy az autokláv bizonyos számú „rekesz”-ből áll, az egyes „rekeszekben” belül a keverés tökéletes és minden egyes rekeszből az anyag a következő rekeszbe áramlik át a beáramlási sebességgel azonos sebességgel. Bár a szóbanforgó rekeszek egymástól a valóságban nincsenek elválasztva, válaszfalaikat csak odaképzeliük, ez a modell a keverés folyamatáról realisabb képet nyújt, mint a tökéletes keverés hipotézise.

Ha a keverés intenzitása i , akkor mivel az S -edrendű gamma-eloszlás relatív szórása, mint láttuk $\frac{1}{\sqrt{S}}$, tehát $i = \frac{1}{S}$, azaz $S = \frac{1}{i}$, tehát a részecskék egy autoklávon való átfutási idejének eloszlása $\frac{1}{i}$ -edrendű gamma-eloszlás.

súnak tekintendő. Ez annyit jelent, hogy amikor r darab egyenként V köbtartalmú autokláv van sorbakapcsolva, amelyekben a keverés $i = \frac{1}{S}$ intenzitású ($S = 2, 3, \dots$), ezt a rendszert gondolatban helyettesítsük rS darab egyenként $\frac{V}{S}$ köbtartalmú olyan autoklávval, amelyben a keverés tökéletes.

Annak indoklására, hogy a (11) képlettel definiált mennyiség valóban adekvát mértékszáma a keverés intenzitásának, ki kell mutatnunk, hogy ésszerű feltevések mellett mindig érvényes a $0 \leq i \leq 1$ egyenlőtlenség. E célból induljunk ki a következő modellből: Az autoklávból időegységenként távozik egy részecske és ugyanakkor belép egy új részecske. Annak a valószínűsége, hogy egy részecske, amely a t időpontban került be az autoklávba, azt a $t + n$ időpontban hagyja el, feltéve, hogy a $t + n - 1$ időpontig nem távozott, legyen P_n . Legyen $C_n = 1 - P_n$. Akkor annak a valószínűsége, hogy egy részecske pontosan N időegységnyi időt töltsön az autoklávban, nyilván $C_1 C_2 \dots C_{N-1} (1 - C_N)$ ($N = 1, 2, \dots$). Jelölje ξ azt az időtartamot, amit a részecske az autoklávban tölt. Be fogjuk bizonyítani, hogy ha $1 \geq C_1 \geq C_2 \geq \dots \geq C_n \geq C_{n+1} \geq \dots$, akkor

$$(12) \quad \frac{D^2(\xi)}{M^2(\xi)} \leq 1.$$

A (12) egyenlőtlenség, figyelembevételével, hogy

$$(13) \quad D^2(\xi) = M(\xi^2) - M^2(\xi)$$

nyilvánvalóan írható az

$$(14) \quad M(\xi^2) \leq 2M^2(\xi)$$

alakba. Nyilvánvalóan

$$(15) \quad M(\xi) = \sum_{n=1}^{\infty} n C_1 C_2 \dots C_{n-1} (1 - C_n) = 1 + \sum_{n=1}^{\infty} C_1 C_2 \dots C_n$$

és

$$(16) \quad M(\xi^2) = \sum_{n=1}^{\infty} n^2 C_1 C_2 \dots C_{n-1} (1 - C_n) = 1 + \sum_{n=1}^{\infty} (2n + 1) C_1 C_2 \dots C_n.$$

Abból a célból, hogy (14)-et bebizonyítsuk, megmutatjuk, hogy érvényes a következő

1. Lemma. Legyen $1 \geq C_1 \geq C_2 \geq \dots \geq C_n \geq \dots$, akkor

$$(17) \quad 1 + \sum_{n=1}^{\infty} (n + 1) C_1 C_2 \dots C_n \leq \left(1 + \sum_{n=1}^{\infty} C_1 C_2 \dots C_n \right)^2.$$

Bizonyítás. Számoljuk meg, hogy ha (17) jobboldalán a négyzetre-emelést elvégezzük, mennyi lesz azon tagok száma, amelyek pontosan n tényezőből állnak. Nyilvánvalóan ilyen tag pontosan $n + 1$ darab lesz; e tagok mindegyike nyilván nagyobb vagy egyenlő, mint $C_1 C_2 \dots C_n$. Így nyerjük a (17) egyenlőtlenséget, ahol egyenlőség akkor és csak akkor áll fenn, ha $C_n = C$ ($n = 1, 2, \dots$), amikor is (17) bal- és jobboldalán egyaránt $\frac{1}{(1-C)^2}$ áll. Mármost (14) könnyen következik lemmánkból, ugyanis (15)-re és (16)-ra való tekintettel (14) a következő alakra hozható:

$$(18) \quad 1 + \sum_{n=1}^{\infty} (2n+1) C_1 \dots C_n \leq 2 \left(1 + \sum_{n=1}^{\infty} C_1 C_2 \dots C_n \right)^2$$

és (17) szerint

$$(19) \quad 1 + \sum_{n=1}^{\infty} (2n+1) C_1 \dots C_n \leq 2 \left(1 + \sum_{n=1}^{\infty} (n+1) C_1 \dots C_n \right) \leq 2 \left(1 + \sum_{n=1}^{\infty} C_1 C_2 \dots C_n \right)^2.$$

Ezzel tehát (14)-et igazoltuk. Az a feltevés, hogy C_n monoton csökkenő sorozat, azt jelenti, hogy minél régebben van egy részecske az autoklávban, annál valószínűbb, hogy azt elhagyja. Ha elfogadjuk, hogy ez a feltétel a valódi keverésnél mindig teljesül (azaz ezt a feltételt belefoglaljuk a valódi keverés definíciójába), akkor tehát bebizonyítottuk, hogy bármely keverés intenzitása 0 és 1 közé eső szám.

A valódi keverési folyamatokat folytonos tárgyalásmód mellett a következőképpen jellemezhetjük. Ha $F(x)$ jelöli egy részecske az autoklávban való átfutási idejének eloszlásfüggvényét és $f(x) = F'(x)$ a sűrűségfüggvényét, akkor közelítőleg $\frac{f(x) \Delta x}{1 - F(x)}$ a feltételes valószínűsége annak, hogy a részecske

x és $x + \Delta x$ időpontok között távozik az autoklávból, feltéve, hogy az x időpontig még nem távozott el. A valódi keverésről akkor beszélünk, ha ez a valószínűség x -nek monoton növekvő függvénye (azaz, ha minél régebben van már a részecske az autoklávban, annál nagyobb a valószínűsége, hogy rövidesen távozik). Ez akkor teljesül, ha $\frac{f(x)}{1 - F(x)} = \frac{d}{dx} \log \frac{1}{1 - F(x)}$ monoton

növekvő, azaz, ha $\log \frac{1}{1 - F(x)}$ konvex függvény.

Nem nehéz megadni a most bebizonyított lemma folytonos analogonját sem. Ezt fejezi ki a

2. Lemma. Legyen $F(x)$ egy pozitív valószínűségi változó eloszlásfüggvénye,

(tehát $F(0) = 0$), tegyük fel, hogy $\log \frac{1}{(1 - F(x))}$ konvex a $(0, +\infty)$ intervallumban. Akkor

$$(20) \quad \int_0^{\infty} x^2 dF(x) \leq 2 \left(\int_0^{\infty} x dF(x) \right)^2.$$

Bizonyítás. Nyilvánvalóan

$$(21) \quad \int_0^{\infty} x^2 dF(x) = 2 \int_0^{\infty} x(1 - F(x)) dx$$

és

$$(22) \quad \int_0^{\infty} x dF(x) = \int_0^{\infty} (1 - F(x)) dx.$$

Legyen $g(x) = -\log(1 - F(x))$, akkor tehát $g(x)$ konvex és a bizonyítandó egyenlőtlenség

$$(23) \quad \int_0^{\infty} x e^{-\ell(x)} dx \leq \left(\int_0^{\infty} e^{-\ell(x)} dx \right)^2.$$

Azonban, mivel $g(x)$ pozitív, monoton növekvő és konvex, és $g(0) = 0$, tehát

$$(24) \quad g(a + b) \geq g(a) + g(b) \quad (a \geq 0, b \geq 0).$$

Mivel

$$(25) \quad \left(\int_0^{\infty} e^{-\ell(x)} dx \right)^2 = \int_0^{\infty} \int_0^{\infty} e^{-\ell(x) - \ell(y)} dx dy = \int_0^{\infty} \left(\int_0^x e^{-\ell(u) - \ell(x-u)} du \right) dx,$$

tehát (24) miatt

$$(26) \quad \left(\int_0^{\infty} e^{-\ell(x)} dx \right)^2 \geq \int_0^{\infty} e^{-\ell(x)} \int_0^x du dx = \int_0^{\infty} x e^{-\ell(x)} dx,$$

ezzel (23)-at és így a 2. lemmát bebizonyítottuk.

A bizonyításból nyilvánvaló, hogy (20)-ban egyenlőség akkor és csak akkor állhat fenn, ha (24)-ben egyenlőség áll fenn minden pozitív a -ra és b -re, tehát ha $g(x)$ lineáris, $g(x) = \lambda x$, ahol $\lambda > 0$, tehát ha

$$F(x) = 1 - e^{-\lambda x}.$$

Más szóval (20)-ban egyenlőség akkor és csak akkor áll fenn, ha $F(x)$ az exponenciális eloszlás eloszlásfüggvénye.

E §. eredményeit a következőképpen foglalhatjuk össze:

Akkor mondjuk, hogy az autoklávban valódi keverés folyik, ha egy részecske az autoklávban való átfutási idejének az eloszlásfüggvényét $F(x)$ -szel jelölve

$\log \frac{1}{1 - F(x)}$ konvex függvény. Ez esetben a keverés intenzitását az átfutási

idő relatív szórásának négyzetével mérjük; a keverés intenzitása 0 és 1 közé eső szám, amely akkor és csak akkor egyenlő eggyel, ha a keverés tökéletes, azaz, ha az átfutási idő exponenciális eloszlással bír.

Azt az esetet, amikor az átfutási idő relatív szórása 1-nél nagyobb, *túlkeverésnek* nevezhetjük. Mint a diszkrét modell tárgyalásából látszik, ez csak akkor fordulhat elő, ha egy részecske távozásának valószínűsége az

autoklávban való tartózkodása során nem növekszik állandóan, hanem időnként csökken. E lehetőség az iparban használt keverő eljárásoknál nyilvánvalóan sohasem fordul elő. Bár elvben e lehetőséget nem lehet kizárni, azonban indokolt, hogy az ilyen folyamatot ne nevezzük valódi keverésnek, hiszen az ilyen folyamat a keveréssel szemben támasztott gyakorlati követelményeknek sem felel meg.

(Beérkezett : 1959. V. 22.)

IRODALOMJEGYZÉK

[1] RÉNYI A.: *Valószínűességszámítás*, Tankönyvkiadó, 1954.

О ПОСЛЕДОВАТЕЛЬНОМ И ПАРАЛЛЕЛЬНОМ СЦЕПЛЕНИИ АВТОКЛАВОВ И О ТЕОРИИ СМЕШЕНИЯ

A. RÉNYI

Резюме

В первых четырех параграфах работы решается следующая проблема поднятая глинозёмным заводом в Almásfűzítő: определить вероятность $F_r(t)$ того, что в r последовательно связанных автоклавах частица бокситной порошки пробудет менее чем t часов, если объем каждого автоклава V , скорость вещества (жидкости, смешенной с бокситным порошком) v и смешение в каждом автоклаве идеально. Ответ таков :

$$(1) \quad F_r(t) = \Gamma_r(\lambda t) ,$$

где $\lambda = \frac{v}{V}$ и $\Gamma_r(x)$ — неполная гамма функция :

$$(2) \quad \Gamma_r(x) = \frac{1}{\Gamma(r)} \int_0^x u^{r-1} e^{-u} du .$$

На основании этого результата исследуется как надо сцеплять автоклавы, чтобы получить наилучший результат, т. е. чтобы $F_r(t_0)$ было наименьшим, где t_0 время, которое каждая частица должна провести в системе для того, чтобы произошла нужная реакция.

В §. 5. изучается проблема неидеального смешения. В автоклаве смешение характеризует функция распределения $F(x)$ времени, которое частица

находится в автоклаве. Доказывается, что в случае действительного смешения $\log \frac{1}{1 - F(x)}$ есть выпуклая функция. Отсюда следует, что

$$(3) \quad \int_0^{\infty} x^2 dF(x) \leq 2 \left(\int_0^{\infty} x dF(x) \right)^2.$$

Иными словами, если обозначит математическое ожидание времени, проведенного частицей в одном автоклаве, через $M(\xi)$ а соответствующую дисперсию через $D(\xi)$, то

$$(4) \quad \frac{D^2(\xi)}{M^2(\xi)} \leq 1,$$

где равенство имеет место в том и только в том случае, если смешение идеальное, т. е. $F(x) = 1 - e^{-\lambda x}$ при $x \geq 0$, где $\lambda > 0$. Величина $\frac{D^2(\xi)}{M^2(\xi)}$ может рассматриваться как *мера интенсивности смешения*.

ON SERIAL AND PARALLEL COUPLING OF AUTOCLAVES AND ON THE THEORY OF MIXING

by

A. RÉNYI

Abstract

In §§. 1—4. the following problem, which has been raised by the Bauxite Works of Almásfüzitő, is solved: to determine the probability $F_r(t)$ that a particle of bauxite powder spends less than t hours in a system consisting of r serially coupled autoclaves, if the volume V of each autoclave, the velocity v of the flow of the material (fluid mixed with the bauxite powder) through the system is given, and it is supposed that the mixing of the material in each autoclave is ideal.

The answer is

$$(1) \quad F_r(t) = \Gamma_r(\lambda t)$$

where $\lambda = \frac{v}{V}$ and $\Gamma_r(x)$ is the incomplete gamma-function

$$(2) \quad \Gamma_r(x) = \frac{1}{\Gamma(r)} \int_0^x u^{r-1} e^{-u} du.$$

On the basis of this result it is discussed how to switch a set of autoclaves to obtain the greatest efficiency, i. e. minimize $F_r(t_0)$ where t_0 is the time

which each particle of the bauxite powder has to spend in the system in order that the necessary chemical reactions should take place.

In § 5. the problem of non-ideal mixing is discussed. The mixing in an autoclave is characterized by the probability distribution function $F(x)$ of the time which a particle spends in the autoclave. It is shown that in case of effective mixing $\log \frac{1}{1 - F(x)}$ is a convex function, and that this implies that

$$(3) \quad \int_0^{\infty} x^2 dF(x) \leq 2 \left(\int_0^{\infty} x dF(x) \right)^2,$$

or by other words, that denoting by $\mathbf{M}(\xi)$ the mean value and by $\mathbf{D}^2(\xi)$ the variance of the time spent by a particle in the autoclave, we have

$$(4) \quad \frac{\mathbf{D}^2(\xi)}{\mathbf{M}^2(\xi)} \leq 1$$

with equality standing if and only if the mixing is ideal, i. e. if $F(x) = 1 - e^{-\lambda x}$ for $x \geq 0$ where $\lambda > 0$. The quantity $\frac{\mathbf{D}^2(\xi)}{\mathbf{M}^2(\xi)}$ may be considered as a *measure of the intensity of mixing*.

ÜBER DIE UMORDNUNG BEDINGT KONVERGENTER REIHEN

von

P. SZÜSZ

Es sei $P = \psi(k)$ ($k = 1, 2, \dots$) eine Permutation der natürlichen Zahlenreihe, d. h. es sei $\psi(k)$ eine für natürliche k definierte Funktion, die nur die natürlichen Zahlen und jede natürliche Zahl genau einmal annimmt. A. S. KRONROD¹ hat die notwendige und hinreichende Bedingung dafür angegeben, dass es eine bedingt konvergente Reihe $\alpha = \sum_{k=1}^{\infty} a_k$ gibt derart, dass die Reihe $P_{\alpha} = \sum_{k=1}^{\infty} a_{\psi(k)}$ divergiert. In der vorliegenden Note zeige ich, dass falls die Glieder der Reihe $\sum_{k=1}^{\infty} a_k$ genügend schnell gegen Null streben, so wird ihre Summe durch die Permutation P nicht beeinträchtigt. (So schnell brauchen die a_k nicht gegen Null zu streben, dass obige Behauptung trivial wird; es braucht nämlich nicht $\sum_{k=1}^{\infty} |a_k| < \infty$ zu sein.) Es wird folgender Satz bewiesen:

Satz. *Es sei $P = \psi(k)$ ($k = 1, 2, \dots$) eine Permutation der natürlichen Zahlen. Dann gibt es eine nur von P abhängige Folge positiver Zahlen c_k ($k = 1, 2, \dots$), die folgendes erfüllt:*

$$(1) \quad c_k \geq c_{k+1} \quad (k = 1, 2, \dots),$$

$$(2) \quad \sum_{k=1}^{\infty} c_k = \infty,$$

so, dass falls $\sum_{k=1}^{\infty} a_k = \alpha$ eine beliebige konvergente numerische Reihe ist, die der Bedingung

$$(3) \quad |a_k| \leq c_k \quad (k = 1, 2, \dots)$$

genügt, so ist die umgeordnete Reihe $\sum_{k=1}^{\infty} a_{\psi(k)}$ auch konvergent, und es ist

$$\sum_{k=1}^{\infty} a_{\psi(k)} = \alpha.$$

¹ On permutation of terms of numerical series, Математический Сборник, Т. 18 (60), 2, 8. 239—276. (Russisch mit englischer Zusammenfassung.)

Bemerkung. Ohne die Beschränkung (2) wäre natürlich alles trivial.

Beweis. Ist $\psi(k)$ gegeben, so definiere man $n(k)$, bzw. $N(k)$ folgendermassen:

$n(k)$ bedeutet die *grösste* Zahl derart, dass unter den

$$\psi(1), \psi(2), \dots, \psi(k)$$

alle Zahlen

$$1, 2, \dots, n(k)$$

vorkommen. Diese Definition ist gleichbedeutend mit

$$n(k) = \min_{l > k} \psi(l) - 1.$$

Da es sich um eine Permutation der natürlichen Zahlenreihe handelt, strebt $n(k)$ mit wachsendem k monoton gegen Unendlich.

$N(k)$ sei definiert durch

$$N(k) = \max_{1 \leq l \leq k} \psi(l).$$

Offenbar gilt

$$(4) \quad N(k) \geq n(k) \quad (k = 1, 2, \dots).$$

Nun sei k_1, k_2, \dots eine streng zunehmende Folge natürlicher Zahlen, die folgendes erfüllt:

$$(5) \quad k_{r+1} - 2k_r + k_{r-1} \geq 0,$$

$$(6) \quad n(k_{r+1}) \geq k_r,$$

$$(7) \quad k_{r+1} \geq N(k_r)$$

(5), (6) und (7) können bei beliebig gegebener Permutation $\psi(k) \dots$ allein durch genügend starkes Anwachsen der k_r erfüllt werden. Man setze ferner

$$(8) \quad c_k = \frac{1}{(k_{r+1} - k_r)r}, \quad \text{falls } k_r \leq k < k_{r+1}.$$

Wegen (5) gilt (1); wegen der Divergenz der harmonischen Reihe ist $\sum_{k=1}^{\infty} c_k$ divergent.

Nun zeige ich, dass bei beliebig klein gegebenem positiven ε

$$(9) \quad \left| \sum_{k=1}^m (a_k - a_{\psi(k)}) \right| < \varepsilon$$

gilt, falls (3) mit den durch (8) definierten c_k gilt und $m \geq m_0(\varepsilon)$ bestet. Es sei r definiert durch

$$(10) \quad k_r \leq m < k_{r+1}.$$

Dann gilt

$$(11) \quad \sum_{k=1}^m (a_k - a_{\psi(k)}) = \sum_{k=n(m)+1}^m a_k - \sum_{\substack{k \leq m \\ \psi(k) \geq n(m)+1}} a_{\psi(k)}.$$

Da wegen (6), (7) und (10) $\psi(k) \leq N(m) \leq N(k_{r+1}) \leq k_{r+2}$ und $n(m) + 1 > n(k_r) > k_{r-1}$ ist, so ist wegen (3)

$$(12) \quad \left| \sum_{k=1}^m (a_k - a_{\psi(k)}) \right| \leq \sum_{k=n(m)+1}^m c_k + \sum_{k=n(m)+1}^{N(m)} c_k$$

also

$$(13) \quad \left| \sum_{k=1}^m (a_k - a_{\psi(k)}) \right| \leq 2 \sum_{k=k_{r-1}+1}^{k_{r+2}} c_k \leq 2 \left(\frac{1}{r-1} + \frac{1}{r} + \frac{1}{r+1} \right).$$

Für genügend grosses m ist die rechte Seite von (13) kleiner als ε . Damit ist (9), also auch unser Satz bewiesen.

(Eingegangen 15. Februar 1959.)

FELTÉTELESEN KONVERGENS SOROK ÁTRENDEZÉSÉRŐL

SZÜSZ P.

Kivonat

Tétel: Legyen $\psi(n)$ ($n = 1, 2, \dots$) a természetes számsor egy permutációja. Akkor megadható egy olyan monoton zérushoz tartó divergens összegű pozitív c_n ($n = 1, 2, \dots$) számsorozat úgy, hogy ha $\sum_{n=1}^{\infty} a_n = A$; és $|a_n| \leq c_n$, akkor $\sum_{n=1}^{\infty} a_{\psi(n)} = A$; másszóval minden permutációhoz megadható oly divergens összegű monoton zérussorozat, hogy ha egy feltételesen konvergens sor tagjai abszolút értékben kisebbek a zérus sorozat tagjainál, akkor a permutáció a sor összegét nem változtatja meg.

О ПЕРЕСТАНОВКЕ ЧЛЕНОВ УСЛОВНО СХОДЯЩИХСЯ РЯДОВ

P. SZÜSZ

Резюме

Теорема. Пусть $\psi(n)$ ($n = 1, 2, \dots$) есть некоторая перестановка последовательности натуральных чисел. Тогда существует такая монотонная сходящаяся к нулю положительная числовая последовательность c_n ($n = 1, 2, \dots$) с расходящейся суммой, для которой имеет место следующее: если $\sum_{n=1}^{\infty} a_n = A$ и $|a_n| \leq c_n$ то $\sum_{n=1}^{\infty} a_{\psi(n)} = A$; иными словами к каждой перестановке можно так подобрать монотонную нулевую последовательность с расходящейся суммой, что, если члены условно сходящегося ряда по модулю меньше членов этой последовательности, то перестановка не изменит суммы ряда.

ON THE EFFICIENCY OF THE SEQUENTIAL PROBABILITY RATIO TEST

by

É. VAS

1. In the following we shall consider the efficiency of the sequential probability ratio test, in case of a simple null hypothesis and a simple alternative hypothesis, relating to the parameter of a distribution of known type. Let us denote by $f(x, \theta)$ the probability density function of the random variable x investigated, by $H_0: \theta = \theta_0$ and $H_1: \theta = \theta_1$ ($\theta_0 < \theta_1$) the null hypothesis and alternative hypothesis, resp., further put

$$z_i = \log \frac{f(x_i, \theta_1)}{f(x_i, \theta_0)}$$

and

$$Z_k = \sum_{i=1}^k z_i.$$

The sequential probability ratio test terminates at the n -th trial, if $\log B < Z_k < \log A$ holds for $k = 1, 2, \dots, n-1$ and either $Z_n \leq \log B$, in which case H_0 is accepted or $Z_n \geq \log A$, in which case H_0 is rejected.

The constants A and B are determined on the base of the risks of errors α and β , as usual, approximately by $A = \frac{1-\beta}{\alpha}$ and $B = \frac{\beta}{1-\alpha}$, resp.

A. WALD and J. WOLFOWITZ [6] proved that in case of two simple hypotheses the sequential probability ratio test is of *optimal character* in the sense that it has the smallest average sample number of all tests with the same strength if either of the hypotheses holds.

A. WALD calculated ([5] p. 57, table 1) for different — practically important — values of α and β the approximate value of the efficiency of the sequential test, relating to the mean of a normally distributed random variable, if θ_0 or θ_1 holds. By efficiency we understand in this case the quotient of the average sample number of the sequential test and the sample number of the current most powerful test, both having the same strength (α, β). If the true value of the parameter is larger than θ_1 or smaller than θ_0 , the sequential test is even more „efficient”. But if the true value of the parameter lies between θ_0 and θ_1 , the calculations of WALD give no information, though in practice the parameter may often be within this interval, and here relatively large sample numbers occur. So the efficiency of the sequential test is of special interest in this case.

A. WALD ([5], p. 172, (A : 77) (A : 78)) determined also upper and lower limits of general validity for the average sample number, for all values of θ . These are the following :

if $\mathbf{M}(\mathbf{z}|\theta) > 0$

then

$$\frac{\bar{L}(\theta) [\log B + \xi'_\theta] + [1 - \bar{L}(\theta)] \log A}{\mathbf{M}(\mathbf{z}|\theta)} \leq \mathbf{M}(\mathbf{n}|\theta) \leq \frac{\underline{L}(\theta) \log B + [1 - \underline{L}(\theta)] [\log A + \xi_\theta]}{\mathbf{M}(\mathbf{z}|\theta)}$$

if $\mathbf{M}(\mathbf{z}|\theta) < 0$

then

$$\frac{\underline{L}(\theta) \log B + [1 - \underline{L}(\theta)] [\log A + \xi_\theta]}{\mathbf{M}(\mathbf{z}|\theta)} \leq \mathbf{M}(\mathbf{n}|\theta) \leq \frac{\bar{L}(\theta) [\log B + \xi'_\theta] + [1 - \bar{L}(\theta)] \log A}{\mathbf{M}(\mathbf{z}|\theta)}$$

where $\bar{L}(\theta)$ and $\underline{L}(\theta)$ denote the upper and lower limit, resp., of the operating characteristic function (given by him too) ([5], pp. 163—164. (A : 31), (A : 34)) and where

$$\xi_\theta = \sup_r \mathbf{M}[\mathbf{z} - r | \theta ; \mathbf{z} \geq r \geq 0]$$

and

$$\xi'_\theta = \inf_r \mathbf{M}[\mathbf{z} + r | \theta ; \mathbf{z} \leq -r \leq 0].$$

WALD ([5], p. 174, (A : 91)) calculated ξ_θ and ξ'_θ for the normal and binomial distribution.

Is $\mathbf{M}(\mathbf{z}|\theta)$ near or equal to 0, the following limits are to be used :

$$\frac{\underline{L}(\theta) [\log B]^2 + [1 - \bar{L}(\theta)] [\log A]^2}{\mathbf{M}(\mathbf{z}^2|\theta)} \leq \mathbf{M}(\mathbf{n}|\theta) \leq \frac{\bar{L}(\theta) [(\log B)^2 + 2 (\log B) \xi'_\theta + \zeta'_\theta] + [1 - \underline{L}(\theta)] [(\log A)^2 + 2 (\log A) \xi_\theta + \zeta_\theta]}{\mathbf{M}(\mathbf{z}^2|\theta)}$$

where

$$\zeta_\theta = \sup_r [(\mathbf{z} - r)^2 | \theta ; \mathbf{z} \geq r \geq 0]$$

and

$$\zeta'_\theta = \sup_r [(\mathbf{z} + r)^2 | \theta ; \mathbf{z} \leq -r \leq 0],$$

ζ_θ and ζ'_θ have been — as far as we know — evaluated for no distribution.

In the following we give for distributions satisfying certain conditions — which are fulfilled among several important distributions by the normal

distribution too — simple and easy calculable expressions for ξ_θ , ξ'_θ , ζ_θ and ζ'_θ ; we calculate with the approximation formula the efficiency of the sequential probability ratio test for the mean of the normal distribution, if the true value of the parameter is $\theta^* = \frac{\theta_0 + \theta_1}{2}$, for several pairs of (α, β) ;

we give further for the normal distribution with the mean value $\theta^* = \frac{\theta_0 + \theta_1}{2}$ more simple and in general better limits, and calculate them for different values of $\alpha = \beta$.

We wish to remark that for 6 concrete cases of binomial distribution H. C. HAMAKER [2] calculated the „efficiency-characteristic function” of the sequential test, while E. PAULSON [3] and S. A. AIVAZIAN [1] considered the problem of the efficiency provided that $\theta_1 \rightarrow \theta_0$.

2. The simpler form of ξ_θ and ξ'_θ is a direct consequence of the following lemma [4]:

Lemma 1. *Let $g(x)$ be either positive for any value of x , or in a finite or infinite interval positive and elsewhere zero, and where $g(x)$ is positive, let there $\log g(x)$ be concave. Let further be $\gamma_1 \leq \gamma_2$ and with the notation*

$$\alpha_i = \frac{\int_{\gamma_i}^{\delta_i} xg(x) dx}{\int_{\gamma_i}^{\delta_i} g(x) dx} \quad (i = 1, 2)$$

let be $\alpha_2 - \alpha_1 \geq \gamma_2 - \gamma_1$ (we regard $\gamma_1 - \gamma_2 = 0$ for $\gamma_1 = \gamma_2 = -\infty$), then for any finite δ_1 the inequality

$$\delta_2 - \delta_1 \geq \gamma_2 - \gamma_1 \quad (\text{i. e. } \alpha_1 - \delta_1 \geq \alpha_2 - \delta_2)$$

holds.

We remark that several important probability density functions — including that of the normal distribution — fulfill the conditions of lemma 1.

Consequently if we denote the density function of the random variable z with parameter value θ by $g(z | \theta)$ and $g(z | \theta)$ fulfills the conditions of lemma 1. for all values of θ , then

$$\mathbf{M}(z - r | \theta; z \geq r \geq 0) = \frac{\int_r^\infty zg(z | \theta) dz}{\int_r^\infty g(z | \theta) dz} - r \quad -\infty < \theta < \infty$$

is a monotonically decreasing function of r and

$$\mathbf{M}(z + r | \theta; z \leq -r \leq 0) = \frac{\int_{-r}^{-\infty} zg(z | \theta) dz}{\int_{-r}^{-\infty} g(z | \theta) dz} + r$$

is a monotonically increasing function of r . Hence we obtain ξ_θ and ξ'_θ at $r = 0$:

$$\xi_\theta = \mathbf{M}(\mathbf{z} | \theta; \mathbf{z} \geq 0)$$

$$\xi'_\theta = \mathbf{M}(\mathbf{z} | \theta; \mathbf{z} \leq 0).$$

Thus the results calculated by WALD directly for the normal distribution may be obtained as a special case in a very simple way (we may assume without loss of generality that $\theta_0 = -\theta_1$ and that the dispersion of the original random variable \mathbf{x} is equal to 1, hence \mathbf{z} — normally distributed too — has $\mathbf{M}(\mathbf{z} | \theta) = 2\theta_1\theta$ and $\mathbf{D}(\mathbf{z} | \theta) = 2\theta_1$):

$$\xi_\theta = 2\theta_1 \left[\theta + \frac{\varphi(\theta)}{\Phi(\theta)} \right]$$

$$\xi'_\theta = 2\theta_1 \left[\theta - \frac{\varphi(\theta)}{\Phi(-\theta)} \right]$$

where

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

and

$$\Phi(x) = \int_{-\infty}^x \varphi(t) dt.$$

3. In determining a simple and easy calculable form of ξ_θ and ξ'_θ we shall make use of the following lemmas 2. and 3.:

Lemma 2. Let us put

$$M_i = \int_a^b \psi(x) g_i(x) dx \quad (i = 1, 2)$$

where $\psi(x) \geq 0$ is monotonically increasing in (a, b) ,
 $g_i(x) \geq 0$ in (a, b)

and

$$\int_a^b g_1(x) dx = \int_a^b g_2(x) dx.$$

If there exists such an $x_0 (a \leq x_0 \leq b)$, that $g_1(x) \leq g_2(x)$ if $a \leq x \leq x_0$ and $g_1(x) \geq g_2(x)$ if $x_0 \leq x \leq b$, then

$$M_1 \geq M_2.$$

Proof.

$$\begin{aligned}
 M_1 - M_2 &= \int_a^b \psi(x) [g_1(x) - g_2(x)] dx = \\
 &= \int_a^{x_0} \psi(x) [g_1(x) - g_2(x)] dx + \int_{x_0}^b \psi(x) [g_1(x) - g_2(x)] dx \geq \\
 &\geq \psi(x_0) \int_a^{x_0} [g_1(x) - g_2(x)] dx + \psi(x_0) \int_{x_0}^b [g_1(x) - g_2(x)] dx = \\
 &= \psi(x_0) \int_a^b [g_1(x) - g_2(x)] dx = 0.
 \end{aligned}$$

Lemma 3. If $g(x)$ fulfills the conditions of lemma 1. then

$$\frac{g(x+h)}{g(x)} = \gamma(x; h)$$

is a monotonically decreasing function of x for any fixed positive value of h .

Proof. Since $\log g(x)$ is concave and all points of a concave arc lie above the chord, for any values $h > 0$ and $h' > 0$ the relation

$$\log g(x+h') + \log g(x+h) \geq \log g(x) + \log g(x+h+h')$$

that is

$$\log g(x+h) - \log g(x) \geq \log g(x+h+h') - \log g(x+h')$$

holds.

Thus $\log g(x+h) - \log g(x) = \log \frac{g(x+h)}{g(x)}$ and consequently $\frac{g(x+h)}{g(x)}$

too are monotonically decreasing functions of x for all positive values of h .

Now we may see that if the probability density function $g(z|\theta)$ fulfills the conditions of lemma 1., then $\mathbf{M}[(\mathbf{z}-r)^2|\theta; \mathbf{z} \geq r \geq 0]$ and $\mathbf{M}[(\mathbf{z}+r)^2|\theta; \mathbf{z} \leq -r \leq 0]$ are monotonically decreasing in r , thus $r=0$ gives the values of ζ_θ and ζ'_θ .

Let namely be

$$M_i = \mathbf{M}[(\mathbf{z}-r_i)^2|\theta; \mathbf{z} \geq r_i \geq 0] = \frac{\int_{r_i}^{\infty} (z-r_i)^2 g(z|\theta) dz}{\int_{r_i}^{\infty} g(z|\theta) dz} \quad (i=1,2).$$

After substitution of $z - r_i = x$ we obtain

$$M_i = \frac{\int_0^{\infty} x^2 g(x + r_i | \theta) dx}{\int_0^{\infty} g(x + r_i | \theta) dx}.$$

Let us denote by

$$g_i(x) = \frac{g(x + r_i | \theta)}{\int_0^{\infty} g(x + r_i | \theta) dx} = c(r_i) g(x + r_i | \theta)$$

and let be $r_2 - r_1 = h > 0$. From lemma 3. it follows that

$$\frac{g_2(x)}{g_1(x)} = \frac{c(r_2)}{c(r_1)} \cdot \frac{g(x + r_2)}{g(x + r_1)}$$

is a monotonically decreasing function of x ; hence, since

$$\int_0^{\infty} g_1(x) dx = \int_0^{\infty} g_2(x) dx = 1,$$

there must exist such an $x_0 \geq 0$ that

$$g_1(x) \leq g_2(x) \text{ for } 0 \leq x \leq x_0$$

and

$$g_1(x) \geq g_2(x) \text{ for } x \geq x_0.$$

$\varphi(x) = x^2$ monotonically increasing in the interval $(0, \infty)$, lemma 2. may be applied, from which it follows that

$$M_1 \geq M_2 \text{ for any values } r_2 \geq r_1 \geq 0.$$

Thus our assertions relative to $\mathbf{M}[(z - r)^2 | \theta; z \geq r \geq 0]$ are proved. Similarly it can be seen that $\mathbf{M}[(z + r)^2 | \theta; z \leq -r \leq 0]$ too is a monotonically decreasing function of r . That is, in the case of distributions fulfilling the conditions of lemma 1.,

$$\zeta_{\theta} = \mathbf{M}[z^2 | \theta; z \geq 0]$$

and

$$\zeta'_0 = \mathbf{M}[z^2 | \theta; z \leq 0].$$

From this follows that for the mean of the normal distribution (with $\sigma = 1$ and $\theta_0 = -\theta_1$)

$$\zeta_{\theta} = 4\theta_1^2 \left[1 + \theta \frac{\varphi(\theta)}{\Phi(\theta)} + \theta^2 \right]$$

and

$$\zeta'_0 = 4\theta_1^2 \left[1 - \theta \frac{\varphi(\theta)}{\Phi(-\theta)} + \theta^2 \right].$$

4. If the true value of the parameter under consideration differs from θ_0 and from θ_1 too, the case when it lies between the two hypothetical values is of special interest; in such cases the sequential probability ratio test may result relatively high sample numbers. Table 1. shows in case of a normal distribution of \mathbf{x} and parameter values $\sigma = 1$, $\theta = \theta^* = \frac{\theta_0 + \theta_1}{2}$ (this means that $\mathbf{M}(\mathbf{z} | \theta^*) = 0$) the efficiency of the sequential test, on base of the approximation formula of WALD, for different values of α and β .

Table 1.

Efficiency of the sequential probability ratio test (approximate values)

$\beta \backslash \alpha$	0,01	0,02	0,03	0,04	0,05
0,01	0,97	0,93	0,90	0,88	0,86
0,02	0,93	0,90	0,87	0,86	0,84
0,03	0,90	0,87	0,85	0,84	0,82
0,04	0,88	0,86	0,84	0,82	0,81
0,05	0,86	0,84	0,82	0,81	0,80

The results obtained by the approximation formula show that the average sample number of the sequential test at risks of errors used in practice does not exceed even for relatively unfavourable parameter values the sample number required by the current test procedure.

The upper and lower limits for the average sample number given by WALD are for the normal distribution and true parameter value θ^* , after substitution of the values of ξ_{θ^*} , ξ'_{θ^*} and ζ_{θ^*} , ζ'_{θ^*} , in case $\alpha = \beta$

$$(1) \quad \frac{(\log A)^3}{4\theta_1^2[\log A + 2\theta_1\varphi(0)]} \leq \mathbf{M}(\mathbf{n} | \theta^*) \leq \frac{(\log A)^3 + 12\varphi(0)\theta_1(\log A)^2 + 4\theta_1^2[1 + 8(\varphi(0))^2]\log A + 16\theta_1^3\varphi(0)}{4\theta_1^2[\log A + 2\theta_1\varphi(0)]}$$

These limits lie — as we shall see in table 3. — rather far from each other.

In case $\alpha = \beta$, i. e. $\log B = -\log A$ from the following relations given by WALD:

$$\mathbf{M}(\mathbf{n} | \theta^*) = \frac{\mathbf{M}(\mathbf{Z}_n^2 | \theta^*)}{\mathbf{M}(\mathbf{z}^2 | \theta^*)}$$

and

$$(\log A)^2 \leq \mathbf{M}(\mathbf{Z}_n^2 | \theta) \leq (\log A)^2 + 2 (\log A) \xi_\theta + \zeta_\theta$$

and from the symmetry of the normal distribution we obtain the following upper and lower limits:

$$(2) \quad \frac{(\log A)^2}{4\theta_1^2} \leq \mathbf{M}(\mathbf{n} | \theta^*) \leq \frac{(\log A)^2}{4\theta_1^2} + \frac{\varphi(0)}{2\theta_1} \log A + 1.$$

This formula gives by very simple calculations limits for $\mathbf{M}(\mathbf{n} | \theta^*)$ and compared with the sample number of the current test procedure, for the efficiency of the sequential process too.

The following table contains the average sample number and the efficiency of the sequential test, for $\alpha = \beta = 0,01; 0,02; 0,03; 0,04; 0,05$ and $\theta_1 = 1; 0,5; 0,25$, calculated by formula (2).

Table 2.

θ_1	$\alpha = \beta$	Average sample number		Efficiency	
		of the sequential probability ratio test			
		lower limit	upper limit	lower limit	upper limit
1	0,01	5,3	7,2	0,98	1,34
	0,02	3,8	5,6	0,89	1,31
	0,03	3,0	4,7	0,86	1,34
	0,04	2,5	4,2	0,82	1,35
	0,05	2,2	3,8	0,80	1,38
0,5	0,01	21,1	23,9	0,98	1,13
	0,02	15,1	17,7	0,89	1,04
	0,03	12,2	14,5	0,86	1,03
	0,04	10,1	12,4	0,82	1,00
	0,05	8,7	10,8	0,80	1,00
0,25	0,01	84,5	89,1	0,98	1,03
	0,02	60,6	64,7	0,89	0,95
	0,03	48,6	52,4	0,86	0,93
	0,04	40,4	43,9	0,82	0,89
	0,05	34,7	38,0	0,80	0,88

As it can be seen from table 2., the upper and lower limits are not too far from one another — and the upper limit of the efficiency even in the most unfavourable cases does not exceed by too much the value 1. We remark the simple fact that in our case of $\alpha = \beta$ the maximum of $\mathbf{M}(\mathbf{n} | \theta)$ lies on behalf of the symmetry of the normal distribution exactly in $\theta = \theta^*$.

In the following table the numerical results obtained by formulas (1) and (2) are compared for different values of $\alpha (= \beta)$ and θ_1 .

Table 3.

Limits of the efficiency of the sequential probability ratio test

θ_1	$\alpha = \beta$	According to formula (1)		According to formula (2)	
		lower limit	upper limit	lower limit	upper limit
1	0,01	0,84	2,12	0,98	1,34
	0,05	0,63	2,47	0,80	1,38
0,5	0,01	0,90	1,48	0,98	1,13
	0,05	0,71	1,49	0,80	1,00
0,25	0,01	0,94	1,22	0,98	1,03
	0,05	0,75	1,13	0,80	0,88

It can be seen that formula (2) which requires very easy calculations, gives also better results.

If we want to calculate the limits of the efficiency of the sequential procedure in case of a normally distributed random variable with mean value $\theta = \theta^*$ but $\alpha \neq \beta$ we can make use of the relation

$$\mathbf{M}\left(\mathbf{n} \mid \theta; \min \frac{1}{A}, B\right) \leq \mathbf{M}(\mathbf{n}, \theta; A, B) \leq \mathbf{M}\left(\mathbf{n} \mid \theta; \max A, \frac{1}{B}\right)$$

and use formula (2), by which we obtain even in this case easier calculable and in general better results.

(Received March 20, 1959.)

REFERENCES

- [1] АЙБАЗАН, С. А.: „Сравнение оптимальных свойств критериев Неймана—Пирсона и Вальда“, *Теория вероятностей и ее применения* IV. 1. 1959.
- [2] НАМАКЕР, А. С.: „The efficiency of sequential sampling for attributes Part. I.“, *Phillips Res.* 8, 1953.
- [3] PAULSON, E.: „A note on the efficiency of the Wald sequential test“, *Ann. Math. Stat.* XVIII, 1947, 447—450.
- [4] SARKADI K.: „Osztályközök megválasztása minták elemeinek osztályba-sorolásánál“, *MTA Alkalmazott Matematikai Intézetének Közleményei*, II.
- [5] WALD, A.: *Sequential Analysis*, New York, Wiley Sons, 1948.
- [6] WALD, A.—WOLFOWITZ, J.: „Optimum character of the sequential probability ratio test“ *Ann. Math. Stat.*, XIX, (1948) 326—339.

A SZEKVENCIAÁLIS VALÓSZÍNŰSÉGI HÁNYADOS PRÓBA EFFICIENCIÁJÁRÓL

VAS É.

Kivonat

Ha az $f(x, \theta)$ sűrűségfüggvényű \mathbf{x} valószínűségi változó esetében a θ paraméterre vonatkozó $H_0: \theta = \theta_0$ ill. $H_1: \theta = \theta_1$ egyszerű hipotézis, ill. ellenhipotézis között akarunk dönteni, adott α és β első-, ill. másodfajú hiba mellett, úgy a szekvenciális valószínűséghányadospróba alapján, mint ismeretes, a

$$\mathbf{Z}_k = \sum_{i=1}^k \mathbf{z}_i = \sum_{i=1}^k \log \frac{f(\mathbf{x}_i, \theta_1)}{f(\mathbf{x}_i, \theta_0)}$$

összeget vizsgáljuk mindaddig, amíg $\log B < \mathbf{Z}_k < \log A$. Ha $k = n$ -re $\mathbf{Z}_n \leq \log B$, úgy a H_0 hipotézist elfogadjuk, míg ha $\mathbf{Z}_n \geq \log A$, úgy a H_0 hipotézist elvetjük. Itt a szokásos közelítés $A = \frac{1-\beta}{\alpha}$ és $B = \frac{\beta}{1-\alpha}$. Wald a szekvenciális eljárás várható mintadarabszámára, $\mathbf{M}(n | \theta)$ -ra, egyrészt közelítő képleteket, másrészt alsó- és felső korlátokat ad. A korlátokban szerepelnek a következő mennyiségek:

$$\xi_\theta = \sup_r \mathbf{M}(\mathbf{z} - r | \theta; \mathbf{z} \geq r \geq 0)$$

$$\xi'_\theta = \inf_r \mathbf{M}(\mathbf{z} + r | \theta; \mathbf{z} \leq -r \leq 0)$$

$$\zeta_\theta = \sup_r \mathbf{M}[(\mathbf{z} - r)^2 | \theta; \mathbf{z} \geq r \geq 0]$$

$$\zeta'_\theta = \sup_r \mathbf{M}[(\mathbf{z} + r)^2 | \theta; \mathbf{z} \leq -r \leq 0].$$

Wald ξ_θ és ξ'_θ értékeit binomiális és normális eloszlás esetében meghatározza. Azon θ^* -hoz közeli paraméterérték esetében, melyre $\mathbf{M}(\mathbf{z} | \theta^*) = 0$, szükséges ζ_θ és ζ'_θ értéke is, ezt azonban nem határozza meg. Jelen dolgozatban a ξ_θ , ξ , ζ_θ , ζ'_θ kifejezésekről bizonyos feltételeknek eleget tevő eloszlások esetében bebizonyítjuk, hogy azok megfelelő szélső értéküket $r = 0$ helyen veszik fel. Így egyszerűbben meghatározhatók normális eloszlás várható értékének vizsgálata esetében a WALD által közvetlenül kiszámított ξ_θ és ξ'_θ értékek, valamint meghatározhatók igen egyszerűen ζ_θ és ζ'_θ értékei is.

Normális eloszlásra, $\theta^* = \frac{\theta_0 + \theta_1}{2}$ várható érték mellett még egyszerűbb és

jobb korlátokat adunk meg; különböző hibanagyságok és hipotézistávolságok mellett kiszámítjuk a szekvenciális eljárás mintanagyságának várható értékét a közelítő képlet segítségével, valamint korlátait és összehasonlítjuk azokat az azonos élességű egyszeres eljáráshoz szükséges mintanagysággal (lásd 1—3. táblázatok).

ОБ ЭФФЕКТИВНОСТИ СЕКВЕНЦИАЛЬНОГО КРИТЕРИЯ ОТНОШЕНИЯ ВЕРОЯТНОСТЕЙ

E. VAS

Резюме

Если в случае случайной величины x с функцией плотности $f(x, \theta)$ нужно выбрать между простой гипотезой и противоположной гипотезой $H_1: \theta = \theta_1$ относительно параметра θ , при данных ошибках α и β первого и второго рода соответственно, то на основании секвенциального критерия отношения вероятностей, как известно, рассматриваем сумму

$$Z_k = \sum_{i=1}^k z_i = \sum_{i=1}^k \log \frac{f(x_i, \theta_1)}{f(x_i, \theta_0)},$$

пока $\log B < Z_k < \log A$. Если при $k = n$ $Z_n \leq \log B$, то принимаем гипотезу H_0 , если же $Z_n \geq \log A$, то не принимаем её. Здесь принято приближение $A = \frac{1-\beta}{\alpha}$ и $B = \frac{\beta}{1-\alpha}$. WALD для ожидаемого числа выборок секвенциального метода $M(n | \theta)$ даёт приближённые формулы и нижние и верхние грани. В гранях фигурируют следующие величины:

$$\xi_\theta = \sup_r M(z - r | \theta; z \geq r \geq 0)$$

$$\xi'_\theta = \inf_r M(z + r | \theta; z \leq -r \leq 0)$$

$$\zeta_\theta = \sup_r M[(z - r)^2 | \theta; z \geq r \geq 0]$$

$$\zeta'_\theta = \sup_r M[(z + r)^2 | \theta; z \leq -r \leq 0].$$

WALD определяет значения ξ_θ и ξ'_θ в случае биномиального и нормального распределения. Для значений параметра, близких к θ^* , для которого $M(z | \theta^*) = 0$, нужны и значения ζ_θ и ζ'_θ , но их он не определяет. В настоящей работе о значениях выражений ξ_θ , ξ'_θ , ζ_θ , ζ'_θ в случае распределений, удовлетворяющих некоторым условиям, доказывается, что они принимают соответствующие экстремальные значения при $r = 0$. Так в случае изучения математического ожидания нормального распределения проще могут быть определены непосредственно вычисленные WALD-ом ξ_θ и ξ'_θ , а также очень просто могут быть вычислены и значения ζ_θ и ζ'_θ . Для нормального распределения при математическом ожидании $\theta^* = \frac{\theta_0 + \theta_1}{2}$ даются ещё более

хорошие и простые грани; при различных величинах ошибок и гипотезных расстояниях с помощью приближённой формулы вычисляется математическое ожидание величины выборки секвенциального метода а также его грани и сравниваются с величиной выборки необходимой для однократного метода аналогичной остроты. (См. таблицы 1—3.)



KÉTPÓLUSÚ ELEKTROMOS HÁLÓZATOKRÓL, IV.

ÁDÁM ANDRÁS

Bevezetés

Cikksorozatunk jelen részében a II. rész vizsgálatait folytatjuk. A II. rész eredményei a VI. osztályú gráfok szerkezetébe nyújtottak bizonyos betekintést; mostani célunk hasonló áttekintést nyerni a IV. osztályú irreducibilis gráfokról. A dolgozat megértéséhez az I. és II. rész ismerete szükséges.

Legyen \mathcal{G} valamely IV. osztályú gráf. \mathcal{G} -nek B típusú élei (ugyanazon kezdőpontra és végpontra vonatkozóan) T_1 tulajdonságú, kétpólusú gráfot alkotnak. Jelöljük ezt a (nyilván kettős él nélküli) gráfot \mathcal{G}' -vel. A II. rész 2a. §-ában bevezetett fogalmakat (\mathcal{G}' alkatrésze, híd, híd határpontja, $\mathfrak{F}(\mathcal{D})$, összetartozó párhuzamos komponensek) ugyanazon értelemben használjuk, mint ott; azonban (\mathcal{G}' már látott definícióján kívül) a híd értelmezése kissé módosul: most a \mathcal{G} gráf C és D típusú éleinek (nem csak a D típusú éleknek!) összességét osztályozzuk az ott látott ekvivalencia-relációval.¹⁾ ²⁾ Ugyancsak módosulni fog az ekvivalens hidak definíciója is.

A VI. osztályú gráfok hídjai — mint az eredményekből kiderült — valóban úgy helyezkedtek el, amint ez a „híd” elnevezés szemléletes tartalma alapján várható volt. Most, mint látni fogjuk, egyéb lehetőségek is előfordulhatnak. Célszerű bevezetnünk a következő definíciókat:

ha $\mathfrak{F}(\mathcal{D})$ párhuzamosan reducibilis gráf és \mathcal{D} -nek van olyan határpontja, amely belső pontja $\mathfrak{F}(\mathcal{D})$ -nek, akkor \mathcal{D} -t *valódi híd*nek nevezzük,

ha $\mathfrak{F}(\mathcal{D})$ sorosan reducibilis, vagy ha $\mathfrak{F}(\mathcal{D})$ párhuzamosan reducibilis

¹⁾ \mathcal{G}' fenti definíciója azért célszerű, mert ha \mathcal{G} -ből csupán a D típusú éleket hagyjuk el, akkor a megmaradó élek a T_1 tulajdonsággal nem rendelkező gráfot alkotnak. — Mint látható, vizsgálatainkban a C és D típusú élek bekapcsolódását fogjuk elemezni a B típusú élek által alkotott gráfba. Közel fekvő probléma volna annak felderítése, hogy a C és D típusú élek *egymáshoz* hogyan kapcsolódhatnak egy IV. osztályú gráf hídjaiban. Nem találtam azonban lehetőséget arra, hogy az utóbbi kérdés vizsgálatánál a kettőnél több pólusú gráfok mélyebb analízise elkerülhető volna; ezt az analízist viszont metodikailag még nem látom időszerűnek.

²⁾ A k_1 és k_2 élek tehát ugyanabba a hídba tartoznak, ha létezik olyan (esetleg zárt) út, amelynek első és utolsó éle k_1 , ill. k_2 , továbbá amely \mathcal{G}' -höz tartozó pontot nem tartalmaz belső pontjaként. — Az ekvivalencia-jellegnek a II. részben (3. lábjegyzet) látott bizonyítása nem teljes, kiegészítendő a következő megfontolással. Megtörténhet, hogy $D \neq E$ esetén az $a[AE] \cdot b[ED]$ (esetleg zárt) út nem létezik; mégpedig pontosan akkor fordul elő ez az eset, ha A belső pontja $b[ED]$ -nek: Ekkor azonban a $k_1^{-1} \cdot b[AD]$ (esetleg zárt) út létezik, és biztosítja, hogy k_1 és k_3 relációban állnak.

és \mathfrak{D} -nek két határpontja van: $\mathfrak{F}(\mathfrak{D})$ kezdőpontja és végpontja, akkor \mathfrak{D} -t *ál-hidnak* nevezzük.³⁾

\mathfrak{G}'' legyen az a részgráfja \mathfrak{G} -nek, amely \mathfrak{G}' éleiből és az ál-hidak éleiből áll.

Egy \mathfrak{D} valódi híd *belső határpontjainak* nevezzük \mathfrak{D} mindazon határpontjait, amelyek $\mathfrak{F}(\mathfrak{D})$ -nek belső pontjai. Nyilvánvaló, hogy \mathfrak{D} -nek $\mathfrak{F}(\mathfrak{D})$ legalább két komponensén van belső határpontja. Két valódi hidat *ekvivalensnek* nevezzünk, ha összes belső határpontjaik megegyeznek. Két ál-hidat *ekvivalensnek* nevezzünk, ha összes határpontjaik megegyeznek.

Az 1. §-ban három segédteét igazolunk, ezek egyike szerint \mathfrak{G}'' T_1 tulajdonságú gráf. A 2. §-ban az ál-hidak bekapcsolódásával fogunk foglalkozni a \mathfrak{G}' gráfba, a 3. §-ban pedig a valódi hidak csatlakozását vizsgáljuk a \mathfrak{G}'' gráfhoz. Mint a 4. §-ban (7. tétel) látni fogjuk, eredményeink a hidak belső szerkezetétől eltekintve a IV. osztályú gráfok teljes leírását szolgáltatják.

A II. és IV. rész szoros kapcsolata lehetővé teszi, hogy a bizonyítások jelentős részében részletes kidolgozás helyett csupán utalhassunk a II. rész állításait igazoló megfontolásokra. A módszerek megegyezésén túlmenően, eredményeink céljukban és részben megfogalmazásukban is megfelelnek a II. rész fő eredményeinek. (A 3., 4., 4a., 7. és 6. tételek formálisan is szembeötlő hasonlóságot mutatnak rendre a II. rész 1., 2., 2a., 3. tételeivel és az 1. tételhez fűzött kiegészítéssel. Tételeink célját tekintve pedig az 1., 3., 4. és részben a 2. tételek feladata szükséges feltételeket adni arra, hogy \mathfrak{G}' ne tartalmazzon kettős élt, mégpedig részben a hidak bekapcsolódási lehetőségeinek szabályozása, részben a belső szerkezetükre kirótt kikötések által,

a 7. tétel feladata megmutatni, hogy a nyert szükséges feltételek elegendőek is, és — amire a II. részben nem volt szükség — az 5. és részben a 2. tétel azt biztosítja, hogy a hidak bármely éle C vagy D típusú él legyen.)

A 2., 4. és 5. tételek látszanak alkalmasaknak arra, hogy a hidak belső szerkezetének későbbi vizsgálatánál kiindulási pontként szerepeljenek.

1. §. Néhány segédétel

1. segédétel. Legyen \mathfrak{G} IV. osztályú gráf, A és B legyenek \mathfrak{G}' tetszőleges pontjai, a és b legyenek \mathfrak{G} -nek olyan pályái, amelyek mind A -n, mind B -n átmennek. Ha a -n A megelőzi B -t, és a csupán \mathfrak{G}' -beli éleket tartalmaz, akkor A megelőzi B -t a b pályán is.⁴⁾

Bizonyítás. Tegyük fel, hogy a tétel feltételei teljesülnek, de a b pályán B előzi meg A -t. Legyen C az első olyan pontja az a pályának, hogy C pontja b -nek is, továbbá az $a[PC]$ és $b[CQ]$ utaknak van közös belső pontjuk. D legyen az $a[PC]$ úton utolsó ilyen közös pont. Ekkor létezik a

$$b[PC] \cdot a^{-1}[CD] \cdot b[DQ]$$

pálya, ellentmondásban azzal, hogy a élei nem kettős élek. ($b[PC]$ és $a[DC]$ C értelmezése miatt idegenek.)

Indukcióval könnyen belátható a következő

³⁾ A II. rész 4. lábjegyzete itt nem érvényes.

⁴⁾ A tétel és bizonyítása a VI. osztályú gráfokra is érvényes.

2. segéd-tétel. Ha A és B egy kettős él nélküli gráf pontjai, akkor (alkalmas módon jelölve egyiküket A_1 -gyel, másikukat A_2 -vel) vannak olyan a és b pályák, hogy az $a[PA_1]$ és $b[A_2Q]$ utak léteznek és idegenek.

3. segéd-tétel. $\mathfrak{G}' T_1$ tulajdonságú gráf.

Bizonyítás. A \mathfrak{G}' -beli éleken nyilván megy át pálya. Legyen k valamely \mathfrak{D} ál-híd éle, menjen át k -n \mathfrak{G} -nek c pályája. A legyen C utolsó \mathfrak{G}' -beli pontja k előtt; B legyen C első \mathfrak{G}' -beli pontja k után. Ekkor $c[AB]$ csupán \mathfrak{D} -beli éleket tartalmaz, és a 2. segéd-tétel miatt alkalmas a és b pályákkal létezik az

$$a[PA_1] \cdot d(A_1A_2) \cdot b[A_2Q]$$

pálya, ahol $d = c[AB]$ vagy $d^{-1} = c[AB]$.

Bizonyítást nem igényel a következő

4. segéd-tétel. Legyen \mathfrak{H} \mathfrak{G}' -nek párhuzamosan reducibilis alkatrésze. \mathfrak{H} (nem feltétlenül különböző) \mathfrak{H}_1 és \mathfrak{H}_2 komponenseinek egy-egy csomópontja legyen A és B , ne teljesüljön a $B < A$ megelőzés. Ha $a(AB)$ belső útja alkalmas hídnak, akkor van \mathfrak{G} -nek olyan b pályája, hogy $a = b[AB]$, továbbá a $b[PA]$ és $b[BQ]$ utak \mathfrak{G}' -beli élekből állnak.

2. §. Az ál-hidak elhelyezkedése

1. tétel. Tekintsük a gráf összes olyan hídjait, hogy $\mathfrak{F}(\mathfrak{D})$ ugyanaz a sorosan reducibilis alkatrésze \mathfrak{G}' -nek. Ekkor teljesülnek a következő **A), B), C)** állítások:

A) \mathfrak{D} bármely határpontja kezdőpontja vagy egyik csomópontja vagy végpontja $\mathfrak{F}(\mathfrak{D})$ -nek.

B) Legyen \mathfrak{D}_α és \mathfrak{D}_β két nem ekvivalens ál-híd. Ekkor a következő négy állítás pontosan egyike igaz:

$\alpha)$ ha A és B tetszőleges határpontjai \mathfrak{D}_α -nak, ill. \mathfrak{D}_β -nak, akkor $A \leq B$,

$\beta)$ \mathfrak{D}_α -nak vannak olyan A_1, A_2 különböző határpontjai, hogy \mathfrak{D}_β -nak tetszőleges B határpontjára $A_1 \leq B \leq A_2$ teljesül, és \mathfrak{D}_α -nak nincs olyan A határpontja, amelyre $A_1 < A < A_2$ volna igaz,

$\gamma)$ és $\delta)$ állítások a \mathfrak{D}_α és \mathfrak{D}_β ál-hidak szerepének felcserélésével nyerhetők az $\alpha)$ és $\beta)$ állításokból.

C) Legyen \mathfrak{D}_α és \mathfrak{D}_β két ekvivalens ál-híd. Ekkor \mathfrak{D}_α és \mathfrak{D}_β határpontjainak száma 2 vagy 3.

Bizonyítás. Az **A)** rész igazolására átvehető a II. rész 1. tétele bizonyításának 2. szakasza.

Nyilvánvaló, hogy a **B)** rész $\alpha)$, $\beta)$, $\gamma)$, $\delta)$ állításai közül legfeljebb egyik teljesül, hiszen bármely ál-hídnak legalább két határpontja van ($\beta)$ és $\delta)$ egyidejű érvényessége azt vonná maga után, hogy \mathfrak{D}_α és \mathfrak{D}_β ekvivalensek); feladatunk bizonyítani, hogy bármely két ál-hídra teljesül valamelyikük. Feltételezzük az ellenkezőt. Ekkor (a két ál-híd jelölését esetleg felcserélve) vannak \mathfrak{D}_α -nak oly A_1, A_2 és \mathfrak{D}_β -nak oly B_1, B_2 határpontjai, hogy $A_1 < B_1 < A_2 < B_2$ igaz.⁵⁾ Ha c az $\mathfrak{F}(\mathfrak{D})$ komponens csomópontjain

⁵⁾ Ez a tény diszkusszióval egyszerűen igazolható, kiindulhatunk pl. úgy, hogy mindkét ál-híd legelső és legutolsó határpontját tekintjük.

átmenő valamely pályája a \mathfrak{G}' gráfnak, továbbá $a(A_1A_2)$ és $b(B_1B_2)$ egy-egy belső útja a \mathfrak{D}_α , illetve \mathfrak{D}_β ál-hidaknak, akkor a

$$c[PA_1] \cdot a \cdot c^{-1}[A_2B_1] \cdot b \cdot c[B_2Q]$$

pálya létezik, tehát $c[B_1A_2]$ kettős út, ami ellentmondás.

Ha **C)** érvénytelen volna, akkor is léteznének az $A_1 < B_1 < A_2 < B_2$ formulát kielégítő határpontok, tehát a látott gondolatmenet a **C)** állítást is igazolja.

Az ál-hidak belső szerkezetére a következő megszorítás érvényes:

2. tétel. Legyen $a(AB)$ belső útja a \mathfrak{D} ál-hídnek, legyenek A és B az $A < B$ feltételnek eleget tevő határpontok. Ekkor van olyan $b(CD)$ belső útja \mathfrak{D} -nek, hogy $C < D$, továbbá alkalmas $k(EF)$ élre $k = a[EF] = b^{-1}[EF]$; és nincs olyan $c(GH)$ belső útja \mathfrak{D} -nek, hogy az $A < G < B < H$ megelőzések igazak, valamint a és c idegenek egymástól. (C , D , G és H \mathfrak{D} határpontjai.)

Megjegyzés. Akkor, ha a \mathfrak{D} ál-hídnek csak két határpontja van, a 2. tétel annyit jelent, hogy \mathfrak{D} bármely pályája tartalmaz kettős élt. Nyilvánvaló, hogy azzal a tulajdonsággal, hogy minden pályája tartalmaz kettős élt, egy irreducibilis gráf pontosan akkor rendelkezik, ha VII. osztályú, egy sorosan reducibilis gráf pontosan akkor rendelkezik, ha legalább egy komponense is ily tulajdonságú,

egy párhuzamosan reducibilis gráf pontosan akkor rendelkezik, ha valamennyi komponense ilyen tulajdonságú.

A 2. tétel **bizonyítása.** A 4. segédtétel szerint van olyan c pályája \mathfrak{G}' -nek, amelynek a része, és amelynek \mathfrak{G}' -be nem tartozó élei pontosan az a út élei. Minthogy c -nek van \mathfrak{G}' -be nem tartozó éle, c másodfajú pálya, és c bármely kettős élet tartalmazza az a út. Legyen k valamely kettős éle c -nek, tekintsünk egy olyan b pályát, amely k -n ellenkező irányban megy át, mint a . A b pályának az a legbővebb $b[CD]$ rész-útja, amely tartalmazza k^{-1} -et, és amelynek belső pontjai \mathfrak{D} -nek is belső pontjai, eleget tesz a tétel első állításának. (A $C < D$ megelőzés az 1. segédtétel indirekt alkalmazásával adódik.)

A tétel második állításának indirekt igazolására legyen d az A , G , B , H pontokon átmenő, \mathfrak{G}' -beli élekből álló pálya. A

$$d[PA] \cdot a \cdot d^{-1}[BG] \cdot c \cdot d[HQ]$$

pálya létezése ellentmondás.

3. §. A valódi hidak elhelyezkedése

3. tétel. Legyen A határpontja a \mathfrak{D} valódi hídnek, A különbözzék $\mathfrak{S} = \mathfrak{F}(\mathfrak{D})$ -nek P^* kezdőpontjától és Q^* végpontjától. Ekkor A csomópontja \mathfrak{S} alkalmas \mathfrak{S}_1 komponensének, és nincs olyan \mathfrak{D}_1 ál-híd, amelyre a következő két állítás mindegyike igaz volna:

$\alpha)$ $\mathfrak{F}(\mathfrak{D}_1) = \mathfrak{S}_1$

$\beta)$ \mathfrak{D}_1 -nek alkalmas B , C határpontjaira igaz a $B < A < C$ megelőzés, továbbá $B \neq P^*$, $C \neq Q^*$ egyike.

Bizonyítás. A tétel első állításának igazolására átvehető a II. rész 1. tétele bizonyításában látott egyik meggondolás: a 4. szakasz második fele. A tétel második állítását könnyű kimutatnunk: ha $b(BC)$ belső útja a \mathfrak{D}_1

ál-hídnak, és $a(AD)$ belső útja a \mathfrak{D} hídnak (ahol D csomópontja \mathfrak{S} valamely \mathfrak{S}_2 ($\neq \mathfrak{S}_1$) komponensének), valamint, ha c és d pályái \mathfrak{S}_1 -nek, illetve \mathfrak{S}_2 -nek, akkor létezik a

$$c[P^*B] \cdot b \cdot c^{-1}[CA] \cdot a \cdot d[DAQ^*]$$

és

$$d[P^*D] \cdot a^{-1} \cdot c^{-1}[AB] \cdot b \cdot c[CQ^*]$$

utak egyike, tehát $c[AC]$ és $c[BA]$ valamelyike kettős út. Ez az ellentmondás igazolja a 3. tételt.

Tekintsük a \mathfrak{G} gráf összes olyan valódi hídjait, amelyekre $\mathfrak{F}(\mathfrak{D})$ ugyanaz $\mathfrak{S}_1, \mathfrak{S}_2, \mathfrak{S}_3$ jelentsék $\mathfrak{S} = \mathfrak{F}(\mathfrak{D})$ különböző komponenseit. Jelöljük $A_k^{(i)}$ -val ($i = 1, 2, 3$) a \mathfrak{S}_i komponens k -ik csomópontját, a csomópontok közé értve a P^* kezdőpontot és Q^* végpontot is. Ha ugyanazon \mathfrak{S}_i komponens több $A_k^{(i)}, A_l^{(i)}, A_m^{(i)}, A_n^{(i)}$ csomópontjáról beszélünk, akkor megállapodunk az $A_k^{(i)} < A_l^{(i)} < A_m^{(i)} < A_n^{(i)}$ rendezésben.

4. tétel. Legyenek $f(AB)$ és $g(CD)$ a (nem feltétlenül különböző) \mathfrak{D}_α illetve \mathfrak{D}_β valódi hidak belső útjai.

1. feltétel. $A = A_k^{(1)}, B = A_n^{(2)}, C = A_m^{(2)}$ és $D = A_l^{(1)}$
2. feltétel. $A = A_k^{(1)}, B = A_m^{(1)}, C = A_l^{(1)}$ és $D = A_n^{(1)}$
3. feltétel. $A = A_k^{(1)}, B = A_m^{(1)}, C = A_l^{(1)}$ és $D = A_n^{(2)}$
4. feltétel. $A = A_k^{(1)}, B = A_m^{(2)}, C = A_l^{(1)}$ és $D = A_n^{(3)}$

Az 1. feltételben megengedjük a $P^* = C$ és $Q^* = B$ egybeesések legfeljebb egyikét⁶⁾, a 2. feltételben pedig a $P^* = A$, és $Q^* = D$ egybeesések bármelyikét (akár mindkettőt is); ettől eltekintve P^* és Q^* különbözni tartoznak A, B, C, D mindegyikétől.

Ha a fenti négy feltétel valamelyike teljesül, akkor f és g nem idegenek egymástól (tehát $\mathfrak{D}_\alpha = \mathfrak{D}_\beta$).

A bizonyítás megegyezik a II. rész 2. tételének bizonyításával. A II. rész 2c §-ának további megfontolásai is átvihetők a IV. osztályú gráfokra. Vezessük be a φ és ψ függvényeket az ott látott módon, de csupán a belső határpontok figyelembevételével. Nevezzük α) állításnak a 4. tételünknek azt az esetét, mikor \mathfrak{D}_α és \mathfrak{D}_β különböző hidak. (Ekkor a 4. tétel azt jelenti, hogy két f, g út az ott felsorolt egyik módon sem helyezkedhet el.) Ekkor igaz a következő

4a tétel. Az α) állítás ekvivalens a következő β) és γ) állításokkal:

β) Legyen A_α belső határpontja a \mathfrak{D}_α valódi hídnak, és A_β belső határpontja a \mathfrak{D}_β valódi hídnak. Legyenek A_α és A_β csomópontjai a \mathfrak{S}_1 komponensnek, legyen érvényes $A_\alpha < A_\beta$. A'_α legyen a \mathfrak{D}_α hídnak valamely belső határpontja a \mathfrak{S}_2 ($\neq \mathfrak{S}_1$) komponensen. Ekkor \mathfrak{D}_β valamely A'_β belső határpontjára igaz az $A_\alpha \leq A'_\beta$ és $A'_\alpha \leq A'_\beta$ megelőzések egyike, továbbá P^* nem határpontja \mathfrak{D}_β -nak és Q^* nem határpontja \mathfrak{D}_α -nak. — Érvényes az az állítás, amely β) eddigi állításából a $<$ és \leq jelek megfordításával, P^* és Q^* felcserélésével kapható.

γ) Mindazok a valódi hidak, amelyekre $\mathfrak{F}^*(\mathfrak{D})$ ugyanaz, teljesítik a következő három állítást:

⁶⁾ Megengedhetnénk egyidejűleg mindkét egybeesést is, azonban ezt az esetet a 2. feltételbe célszerűbb beiktatni.

(1) Ha \mathfrak{D}_α és \mathfrak{D}_β ekvivalens hidak és belső határpontjaik száma legalább 4, akkor ezek a határpontok páronként különböző komponenseken vannak. Ha \mathfrak{D}_α és \mathfrak{D}_β ekvivalens hidak 3 belső határpontja \mathfrak{H} két komponensen helyezkedik el, akkor mind \mathfrak{D}_α -nak, mind \mathfrak{D}_β -nek minden határpontja belső.

(2) Legyen a \mathfrak{D}_α és \mathfrak{D}_β nem ekvivalens hidak két alkalmas A_α, A_β belső határpontjára $A_\alpha < A_\beta$. Ha A_α° és A_β° belső határpontjai \mathfrak{D}_α -nak, ill. \mathfrak{D}_β -nak, és ugyanazon a komponensen helyezkednek el, akkor $A_\alpha^\circ \leq A_\beta^\circ$; továbbá P^* nem határpontja \mathfrak{D}_β -nak és Q^* nem határpontja \mathfrak{D}_α -nak.

(3) Ha $\mathfrak{F}^*(\mathfrak{D})$ legalább három komponensből áll, továbbá a \mathfrak{H}_1 komponensére és a különböző $\mathfrak{D}_\alpha, \mathfrak{D}_\beta$ hidakra

$$\psi(\mathfrak{D}_\alpha, \mathfrak{H}_1) = 1, \psi(\mathfrak{D}_\beta, \mathfrak{H}_1) = 1 \text{ és } \varphi(\mathfrak{H}_1) > 1$$

teljesülnek, akkor a $\psi(\mathfrak{D}, \mathfrak{H}_1) = 1$ feltételt kielégítő hidaknak \mathfrak{H}_1 -nek kívüli belső határpontjai egyetlen \mathfrak{H}_2 komponensen vannak, és $\varphi(\mathfrak{H}_2) = 1$.

Bizonyítás gyanánt a II. rész 2a tételének igazolása szolgálhat;⁷⁾ kiegészítésül csak annyit kell megjegyeznünk, hogy a β) állítás P^* -ra és Q^* -ra vonatkozó részének hamissága a 4. tétel 1. esetével idézne elő ellentmondást.

5. tétel. Legyen a (AB) belső útja valamely \mathfrak{D} valódi hídnek, ahol $A = P^*$, $B = Q^*$ valamelyike igaz, A és B \mathfrak{D} határpontjai. (P^* és Q^* $\mathfrak{F}(\mathfrak{D})$ kezdőpontja és végpontja.) Ekkor a -nak van olyan k éle, hogy k^{-1} része \mathfrak{D} alkalmas $b(CD)$ belső útjának, ahol $C \neq Q^*$, $D \neq P^*$ és nem teljesül $D < C$.

Bizonyítás. Az a út (csupán \mathfrak{G}' -beli élekkel) kiegészíthető \mathfrak{G} egy pályájává (4. segédtétel), tehát tartalmaz kettős élt. Legyen egy kettős éle k ; a k élen fordított irányban áthaladó bármely pályának az a részútja, amely tartalmazza k -t, valamint \mathfrak{D} -nek határpontjait köti össze és belső út \mathfrak{D} -ben, eleget tesz a tétel állításának (vö. 1. segédtétel).

4. §. Általános áttekintés

6. tétel. Nincs olyan ál-híd, amelynek pontosan a \mathfrak{G} gráf P kezdőpontja és Q végpontja a határpontjai. Ha \mathfrak{G}' párhuzamosan reducibilis gráf, akkor \mathfrak{G}' bármely két komponense összetartozó. Ha \mathfrak{G}' sorosan reducibilis gráf, akkor van olyan ál-híd, amelynek határpontjai között P is, Q is fellépnek. Igaz a következő két állítás valamelyike:

a) van legalább egy ál-híd \mathfrak{G} -ben,

β) van \mathfrak{G} -nek olyan \mathfrak{D} valódi hidja, hogy \mathfrak{D} határpontjai között az $\mathfrak{F}(\mathfrak{D})$ alkatrész kezdőpontja vagy végpontja fellép.

Bizonyítás. A tétel első két állításának érvénytelensége esetén \mathfrak{G} párhuzamosan reducibilis gráf volna.⁸⁾ Legyen \mathfrak{G}' sorosan reducibilis. Ha P nem lépne fel az $f(\mathfrak{D}) = \mathfrak{G}'$ feltételt teljesítő egyik \mathfrak{D} ál-híd határpontjaként sem, akkor \mathfrak{G}' első komponense \mathfrak{G} -nek is komponense volna, ami ellentmondás. Legyen A a \mathfrak{G}' gráf utolsó olyan csomópontja (Q -t is ideértve), hogy P és A összeköthetőek alkalmas ál-híd belső útjával. Feltételezzük, hogy $A \neq Q$. Ekkor van olyan a pályája \mathfrak{G} -nek, hogy a nem megy át A -n, és a -nak nyilván van olyan $a[BC]$ részútja, amely belső útja valamely (\mathfrak{D} -től különböző) ál-hídnek, és amelyben B és C \mathfrak{G}' -nek a $B < A < C$ fel-

⁷⁾ A II. rész 2. tételét természetesen mindenütt e rész 4. tétele pótolja, határpont helyett belső határpont értendő.

⁸⁾ Vö. a II. rész 1. tétele bizonyításának 5. szakaszával.

tételt teljesítő csomópontjai. ($P \neq B$, de esetleg $Q = C$). Ez azonban ellentmond az 1. tétel B) állításának. A tétel utolsó mondatát az teszi világossá, hogy a gráfnak (a IV. osztály értelmezése miatt) van C típusú éle; ha pedig egy híd tartalmaz C típusú élt, akkor az α) vagy β) állításban leírt módon kell elhelyezkednie.

A következő tételben a \mathfrak{G} gráftól csupán annyit kívánunk meg, hogy bármely pontjából induljon ki él.

7. tétel. *Legyen adva egy \mathfrak{G} gráf. \mathfrak{G}' legyen olyan részgráfja \mathfrak{G} -nek, hogy \mathfrak{G}' kezdőpontja és végpontja megegyezik \mathfrak{G} kezdőpontjával, ill. végpontjával, továbbá \mathfrak{G}' felépíthető $P \circ \dots \circ Q$ alapelemből soros és párhuzamos kapcsolásokkal. \mathfrak{G} -nek \mathfrak{G}' -be nem tartozó éleit osztályozzuk H -osztályokba ugyanazzal az eljárással, amellyel az ál-hidakat és valódi hidakat értelmeztük. Legyenek ezek a H -osztályok T_3 tulajdonságú n -gráfok.⁹⁾ Ha az 1., 2., 3., 4., 5. és 6. tételek érvényben maradnak, akkor \mathfrak{G} IV. osztályú irreducibilis gráf, amelynek pontosan a \mathfrak{G}' -be nem tartozó élei C vagy D típusúak.*

Bizonyítás. A 2. segédtétel következtében \mathfrak{G} T_1 tulajdonságú. A 6. tétel biztosítja, hogy \mathfrak{G} irreducibilis gráf, amely tartalmaz C típusú élt. A II. rész 3. tétele bizonyításának gondolatmenetével igazolható, hogy \mathfrak{G}' nem tartalmaz kettős élt (tehát C típusú élt sem). A 2. esetben most a következő állítások egyikével adódna ellentmondás: az 1. tétel B), C) részei, a 2. vagy 3. tétel második része, a 4. tétel.)

Ha \mathfrak{G} valamely a pályája tartalmaz \mathfrak{G}' -n kívüli élt, akkor ez a pálya másodfajú. Van ugyanis olyan $a[AB]$ út, amely belső útja valamely \mathfrak{D} hídnak, A és B a híd határpontjai. Négy eset lehetséges, amelyeket alább részletezünk; mind a négy esetben a tekintett $a^{-1}[BA]$ út, vagy b és b^{-1} egyike kiegészíthető \mathfrak{G} valamely c pályájává, c és $a[AB]$ alkalmas élen ellenkező irányban futnak át; így tehát a H -osztályok minden éle C vagy D típusú.

1. eset. Ha A és B határpontjai az $\mathfrak{F}(\mathfrak{D})$ párhuzamosan redubicilis gráf különböző komponenseinek, akkor tekintsük az $a^{-1}[BA]$ utat.

2. eset. Ha \mathfrak{D} ál-híd, akkor tekintsük a 2. tétel első részében említett b utat.

3. eset. Ha \mathfrak{D} valódi híd, A és B pedig $\mathfrak{F}(\mathfrak{D})$ kezdőpontja és végpontja, akkor tekintsük az 5. tételben említett b utat.

4. eset. \mathfrak{D} valódi híd, A és B csomópontjai ugyanazon komponensének, megengedve az $A = P^*$, $B = Q^*$ egybeesések legfeljebb egyikét (ahol P^* és Q^* kezdőpontja, ill. végpontja). Ekkor k legyen olyan éle $a[AB]$ -nek, amely $\mathfrak{F}(\mathfrak{D})$ szóbanforgó komponense valamely (P^* -től és Q^* -től különböző) C csomópontjához illeszkedik (C tehát megegyezik A -val vagy B -vel), és $b(CD)$ legyen olyan belső útja \mathfrak{D} -nek, amely átmegy k -n, és amelyre a D pont $\mathfrak{F}(\mathfrak{D})$ -nek más komponensében van, mint C .

(Beérkezett: 1959. I. 19.)

IRODALOM

ÁDÁM A.: „Kétpólusú elektromos hálózatokról I—II.” *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* I: 2 (1957) 211—218. II: 3 (1958) 67—79.

⁹⁾ Lásd: II. rész, 69. oldal.

ÜBER ZWEIPOLIGE ELEKTRISCHE NETZE, IV.

A. ÁDÁM

Zusammenfassung

In diesem Teile werden die Untersuchungen des Teiles II. fortgesetzt. Wir studieren die Graphen von Klasse IV. (laut der Terminologie, die in der Einleitung des Teiles II. gegeben war), d. h. die (durch Reihen- und Parallelschaltung) irreduziblen Graphen mit den folgenden zwei Eigenschaften:

der Graph enthält (mindestens) eine Bahn, die keine doppelte Kante hat, und

der Graph enthält eine solche nicht doppelte Kante k , dass jede Bahn des Graphen, die k durchläuft, eine doppelte Kante enthält.

Die Untersuchungen haben eine enge Analogie mit dem Teile II., aber die Ergebnisse sind komplizierter. Die Kanten derjenigen Bahnen, die keine doppelte Kante enthalten, bilden den Graphen \mathfrak{G}' . Die übrigen Kanten werden auch hier in Brücken zerteilt, und es werden die möglichen Anlagen der Brücken in demselben Sinne festgesetzt, wie im Teile II. Dort die Benennung "Brücke" wurde von den Resultaten gerechtfertigt; jetzt ist das nicht wahr in allgemeinen, und es ist zweckmässig echte Brücken (studiert im § 3.) und Pseudobrücken (studiert im § 2.) zu unterscheiden.

О ДВУХПОЛЮСНЫХ ЭЛЕКТРИЧЕСКИХ СЕТЯХ, IV.

A. ÁDÁM

Резюме

В настоящем сообщении продолжаются исследования части II. Изучаются графы класса IV (согласно терминологии введения к части II) т. е. те (последовательно и параллельно) неприводимые графы, которые удовлетворяют двум следующим условиям:

граф содержит путь, не имеющий двойных рёбер,

граф содержит такое простое ребро k , что проходящий через k любой путь содержит двойное ребро.

Исследования находятся в тесной аналогии с частью II, но результаты более сложные. Рёбра путей без двойных рёбер образуют граф \mathfrak{G}' . Остальные рёбра и здесь причисляются к мостам и определяются возможные расположения мостов в том же смысле, как в части II. Там название «мост» оправдывали результаты; сейчас это, вообще говоря, не имеет места и целесообразно различать настоящие мосты (изучаемые в § 3) и лжемосты (изучаемые в § 2).

ANWENDUNG DER OPERATORENRECHNUNG VON MIKUSIŃSKI ZUR LÖSUNG GEWÖHNLICHER DIFFERENTIALDIFFERENZENGLEICHUNGEN

von
T. FÉNYES

§ 1. Einleitende Bemerkungen

Betrachten wir die folgende Differentialdifferenzengleichung

$$(1) \quad \sum_{\mu=0}^m \sum_{v=0}^n a_{\mu v} x^{(\mu)}(t + v\alpha) = g(t), \quad 0 \leq t < \infty$$

wo $g(t)$ eine in dem betrachteten Intervall stetige Funktion ist. Ist $\alpha > 0$ so setzen wir voraus, dass die Koeffizienten $a_{\mu v}$, $v = 0, 1, \dots, n-1$ alle gleich Null sind. Das bedeutet, dass die höchste Ableitung nur mit dem Argument von maximaler Verschiebung auftreten soll. Wir werden die betrachtete Differentialdifferenzengleichung mit Hilfe der Operatorenrechnung lösen, und zeigen, dass die zu der Gleichung (1) gehörenden Anfangsbedingungen sich automatisch aus der Operatorgleichung ergeben.

§ 2. Darstellung der Lösung

Es sei

$$\{x(t)\} = f(s)$$

ein Element des MikusiŃskischen Quotientenkörpers. Mit Hilfe des Verschiebungsoperators und des Differentiationsoperators lässt sich $x^{(\mu)}(t + v\alpha)$ einfach darstellen. Aus den Regeln der Operatorenrechnung folgt es, dass

$$(2) \quad e^{-\alpha v s} \{x(t + v\alpha)\} = \begin{cases} \{x(t)\} & \text{für } \alpha v \leq t < \infty \\ 0 & \text{für } 0 \leq t < \alpha v. \end{cases}$$

ist.

Wir wollen $x(t)$ auf der ganzen positiven Halbgeraden in der Operatorschreibweise darstellen. Wenden wir das Integral von Ryll-Nardzewski an, so erhalten wir

$$(3) \quad \int_0^{\alpha v} x(\lambda) e^{-\lambda s} d\lambda = \begin{cases} \{x(t)\} & \text{für } 0 \leq t \leq \alpha v \\ 0 & \text{für } t > \alpha v. \end{cases}$$

Aus (2) und (3) folgt, dass

$$(4) \quad \{x(t + \alpha v)\} = e^{\alpha v s} \left[f(s) - \int_0^{\alpha v} x(\lambda) e^{-\lambda s} d\lambda \right]$$

ist.

Ziehen wir in Betracht, dass

$$x^{(\mu)}(t + v\alpha) = s^\mu x(t + v\alpha) - s^{\mu-1}x(v\alpha) - s^{\mu-2}x'(v\alpha) - \dots - x^{(\mu-1)}(v\alpha)$$

$$(5) \quad \begin{aligned} 0 &\leq \mu \leq m, \\ 0 &\leq v \leq n \end{aligned}$$

ist. Wenn wir (4) und (5) in (1) substituieren, erhalten wir die folgende gewöhnliche algebraische Gleichung:

$$(6) \quad \sum_{\mu=0}^m \sum_{v=0}^n a_{\mu v} \left\{ s^\mu e^{\alpha v s} \left[f(s) - \int_0^{\alpha v} x(\lambda) e^{-\lambda s} d\lambda \right] - \right. \\ \left. - x(\alpha v) s^{\mu-1} - x'(\alpha v) s^{\mu-2} - \dots - x^{(\mu-1)}(\alpha v) \right\} = \{g(t)\}.$$

Aus dieser Operatorgleichung erhalten wir die allgemeine Lösung unserer Differentialdifferenzgleichung. Die Lösung lautet:

$$(7) \quad \begin{aligned} f(s) = \{x(t)\} = & \left[\{g(t)\} + \sum_{\mu=0}^m \sum_{v=0}^n a_{\mu v} s^\mu e^{\alpha v s} \int_0^{\alpha v} x(\lambda) e^{-\lambda s} d\lambda + \right. \\ & + \sum_{\mu=0}^m \left\{ s^{\mu-1} \sum_{v=0}^n a_{\mu v} x(\alpha v) + s^{\mu-2} \sum_{v=0}^n a_{\mu v} x'(\alpha v) + \dots \right\} : \\ & \left. : \left[\sum_{v=0}^n e^{\alpha v s} \sum_{\mu=0}^m a_{\mu v} s^\mu \right] \right]. \end{aligned}$$

Dieser Ausdruck zeigt uns, dass man $x(t)$ in $0 \leq t \leq \alpha n$ als $m-1$ -mal differenzierbare Funktion angeben muss. Damit ist die Lösung eindeutig bestimmt.

Wir haben unser Problem schon gelöst. Wollen wir (7) in Symbolen der klassischen Analysis schreiben, so müssen wir den Ausdruck auf folgende Weise in eine unendliche Reihe entwickeln.

Es ergibt sich durch eine einfache Umformung:

$$(8) \quad \left(\sum_{v=0}^n e^{\alpha v s} \sum_{\mu=0}^m a_{\mu v} s^\mu \right)^{-1} = \frac{e^{-\alpha n s}}{\sum_{\mu=0}^m a_{\mu n} s^\mu \left(1 + \frac{\sum_{v=0}^{n-1} e^{\alpha v s} \sum_{\mu=0}^m a_{\mu v} s^\mu}{\sum_{\mu=0}^m a_{\mu n} s^\mu} e^{-n\alpha s} \right)}.$$

Der Operator

$$e^{-nas} \frac{\sum_{v=0}^{n-1} e^{avs} \sum_{\mu=0}^m a_{\mu v} s^\mu}{\sum_{\mu=0}^m a_{\mu n} s^\mu}$$

stellt, wie sich das leicht einsehen lässt, eine Funktion aus der Klasse K dar. Wir können daher den Operator

$$\left(1 + e^{-nas} \frac{\sum_{v=0}^{n-1} e^{avs} \sum_{\mu=0}^m a_{\mu v} s^\mu}{\sum_{\mu=0}^m a_{\mu n} s^\mu} \right)^{-1}$$

in eine Reihe entwickeln, welche im Sinne der Operatorenrechnung konvergent ist :

$$(9) \quad \left(\dots \dots \right)^{-1} = \sum_{k=0}^{\infty} (-1)^k \left(\frac{\sum_{v=0}^{n-1} e^{avs} \sum_{\mu=0}^m a_{\mu v} s^\mu}{\sum_{\mu=0}^m a_{\mu n} s^\mu} \right)^k e^{-nkas}.$$

Setzen wir (9) in die Gleichung (8) ein, so erhalten wir für die Lösung (7) den folgenden Ausdruck :

$$(10) \quad f(s) = \{x(t)\} = \left\{ \{g(t)\} + \sum_{\mu=0}^m \sum_{v=0}^n a_{\mu v} s^\mu e^{avs} \int_0^{av} x(\lambda) e^{-\lambda s} d\lambda + \right. \\ \left. + \sum_{\mu=0}^m \left[s^{\mu-1} \sum_{v=0}^n a_{\mu v} x(av) + s^{\mu-2} \sum_{v=0}^n a_{\mu v} x'(av) + \dots \sum_{v=0}^n a_{\mu v} x^{(\mu-1)}(av) \right] \right\} \times \\ \times \frac{e^{-nas}}{\sum_{\mu=0}^m a_{\mu n} s^\mu} \sum_{k=0}^{\infty} (-1)^k \left(\frac{\sum_{v=0}^{n-1} e^{avs} \sum_{\mu=0}^m a_{\mu v} s^\mu}{\sum_{\mu=0}^m a_{\mu n} s^\mu} \right)^k e^{-nkas}.$$

Aus diesem Ausdruck kann man $x(t)$ in der üblichen Schreibweise aufschreiben. Sind m und n grosse Zahlen, so ergibt sich zwar eine ziemlich komplizierte Formel, prinzipielle Schwierigkeiten treten jedoch nicht auf.

Sind die Koeffizienten a_{mv} , $v = 0, 1 \dots n-1$ nicht alle gleich Null, und ist $n > 1$ so ist in der Operatorenrechnung kein ähnlicher Entwicklungssatz für den Operator (9) bekannt. Ist $n = 1$ so gilt die folgende Entwicklung :

$$\left(1 + \frac{\sum_{\mu=0}^m a_{\mu 0} s^\mu}{\sum_{\mu=0}^m a_{\mu 1} s^\mu} e^{-a} \right)^{-1} = \sum_{k=0}^{\infty} (-1)^k \left(\frac{\sum_{\mu=0}^m a_{\mu 0} s^\mu}{\sum_{\mu=0}^m a_{\mu 1} s^\mu} \right)^k e^{-kas}.$$

Wir werden $f(s)$ in diesem Spezialfall nicht näher betrachten. Ist $\alpha < 0$ so setzen wir voraus dass die Koeffizienten a_{nv} , $v = 1, 2 \dots n$ alle gleich Null sind. Das bedeutet, dass die höchste Ableitung nur mit dem Argument ohne Verschiebung auftreten soll. Wir erhalten die folgende Lösung:

$$(11) \quad \{x(t)\} = \left[g(t) - \sum_{\mu=0}^m \sum_{v=0}^n a_{\mu v} s^{\mu} \int_0^{\beta v} x(\lambda - v\beta) e^{-\lambda s} d\lambda + \right. \\ \left. + \sum_{\mu=0}^m \left\{ s^{\mu-1} \sum_{v=0}^n a_{\mu v} x(-v\beta) + \dots + \sum_{v=0}^n a_{\mu v} x^{(\mu-1)}(-v\beta) \right\} \right] \times \\ \times \frac{1}{\sum_{\mu=0}^m a_{\mu 0} s^{\mu}} \sum_{k=0}^{\infty} (-1)^k \left(\frac{\sum_{v=1}^n e^{-v\beta} \sum_{\mu=0}^m a_{\mu v} s^{\mu}}{\sum_{\mu=0}^m a_{\mu 0} s^{\mu}} \right)^k$$

wo $\beta = -\alpha$ ist. Ist $x(t)$ in $-\beta n \leq t \leq 0$ als eine $m-1$ -mal differenzierbare Funktion gegeben, so ist die Lösung eindeutig bestimmt.

§ 3. Ein Beispiel

Wir werden die Lösungsmethode an Hand eines Beispiels schildern. Lösen wir die folgende Differentialdifferenzgleichung:

$$(12) \quad x''(t + \pi) + 2x'(t) + x(t) = e^t$$

mit der Anfangsbedingung

$$x(t) = \sin t \quad \text{für} \quad 0 \leq t \leq \pi.$$

Es sind

$$(13) \quad \begin{aligned} x(0) &= 0 \\ x(\pi) &= 0 \\ x'(0) &= 1 \\ x'(\pi) &= -1 \end{aligned}$$

und

$$(14) \quad \int_0^{\pi} \sin \lambda e^{-\lambda s} d\lambda = \frac{1 + e^{-\pi s}}{1 + s^2}.$$

Weiterhin ist

$$(15) \quad \begin{aligned} \{x''(t + \pi)\} &= s\{x'(t + \pi)\} - x'(\pi) = s^2\{x(t + \pi)\} - sx(\pi) - x'(\pi) = \\ &= s^2 e^{\pi s} f(s) - s^2 \frac{1 + e^{\pi s}}{1 + s^2} + 1. \end{aligned}$$

Mit (15) und

$$\{x'(t)\} = sf(s) - x(0) = sf(s)$$

erhalten wir die folgende Operatorgleichung

$$s^2 e^{\pi s} f(s) - s^2 \frac{1 + e^{\pi s}}{1 + s^2} + 1 + 2 s f(s) + f(s) = \{e^{t^2}\}$$

woraus

$$(16) \quad f(s) = \{x(t)\} = \frac{\{e^{t^2}\} + s^2 \frac{1 + e^{\pi s}}{1 + s^2} - 1}{s^2 e^{\pi s} + 2 s + 1}$$

folgt. Mit Anwendung der oben betrachteten Reihenentwicklung ergibt sich

$$(17) \quad (s^2 e^{\pi s} + 2 s + 1)^{-1} = \frac{e^{-\pi s}}{s^2 + (1 + 2 s) e^{-\pi s}} = \frac{e^{-\pi s}}{s^2} \cdot \frac{1}{1 + \left(\frac{1 + 2 s}{s^2}\right) e^{-\pi s}} =$$

$$= \frac{e^{-\pi s}}{s^2} \sum_{k=0}^{\infty} (-1)^k \left(\frac{1 + 2 s}{s^2}\right)^k e^{-k\pi s}.$$

Setzen wir nun (17) in (16), so erhalten wir für die Lösung

$$(18) \quad \{x(t)\} = f(s) = e^{-\pi s} \left(\frac{\{e^{t^2}\}}{s^2} + \frac{1 + e^{\pi s}}{1 + s^2} - \frac{1}{s^2} \right) \sum_{k=0}^{\infty} (-1)^k \left(\frac{1 + 2 s}{s^2}\right)^k e^{-k\pi s}.$$

Aus diesem Ausdruck kann man die Lösung ohne besonderen Schwierigkeiten, in der üblichen Schreibweise als eine Funktion von t aufschreiben.

(Eingegangen 20. Mai, 1959.)

LITERATURVERZEICHNIS

[1] MIKUSIŃSKI, J.: *Operatorenrechnung*, VEB Deutscher Verlag der Wissenschaften, Berlin, 1957.

A MIKUSIŃSKI-FÉLE OPERÁTORSZÁMÍTÁS ALKALMAZÁSA ÁLLANDÓ EGYÜTTHATÓS RETARDÁLT DIFFERENCIÁLEGYENLETEK MEGOLDÁSÁRA

FÉNYES T.

Kivonat

Szerző a Mikusinski-féle operátorszámítás segítségével előállítja a

$$\sum_{\mu=0}^m \sum_{\nu=0}^n a_{\mu\nu} x^{(\mu)}(t + \nu\alpha) = g(t)$$

difference-difference equation operátormegoldását.

Megmutatja, hogy a megoldás egyértelműségét biztosító, megadandó kezdeti feltételek a módszer alkalmazása révén az operátoregyenletből leolvashatók. Ezután egy eljárást mutat meg, amelynek alkalmazásával a megoldás a klasszikus analízis jelöléseivel is felírható.

ПРИМЕНЕНИЕ ОПЕРАЦИОННОГО ИСЧИСЛЕНИЯ К РЕШЕНИЮ ОБЫКНОВЕННЫХ РАЗНОСТНО-ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

T. FÉNYES

Резюме

Автор с помощью операционного исчисления от Mikusiński находит операционное решение разностно-дифференциального уравнения

$$\sum_{\mu=0}^m \sum_{\nu=0}^n a_{\mu\nu} x^{(\mu)}(t + \nu\alpha) = g(t) .$$

Доказывает, что начальные условия, обеспечивающие единственность решения, могут быть определены с помощью операторного уравнения. После этого даётся метод, с помощью которого решения может быть записано и обозначениями классического анализа.

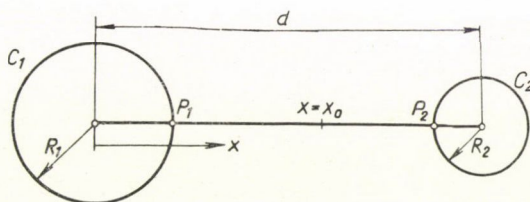
UNE REMARQUE SUR LA COMMUNICATION INTERCOSMIQUE

par

GEORGES ADLER

Le but de cet article est de chercher le temps minimum suffisant pour atteindre un corps céleste après avoir donné l'accélération maximum à laquelle l'organisme humain peut être soumis. Dans ce qui suit, nous ferons abstraction des problèmes techniques.

Considérons deux corps célestes C_1 et C_2 de formes sphériques, de masses respectives M_1 et M_2 ayant une distance fixe d l'un de l'autre. Supposons que leurs centres de gravité se trouvent aux centres des sphères, de plus, pour simplifier la situation, supposons qu'ils ne se meuvent pas (voir la figure).



Figure

En premier lieu, nous examinons le problème élémentaire suivant :

Problème 1 : On part du point P_1 avec une vitesse nulle vers le point P_2 en suivant l'axe x . On veut progresser premièrement en accélérant jusqu'au point $x = x_0$, puis en ralentissant jusqu'au point P_2 de telle façon que la valeur absolue de l'accélération soit toujours égale à une constante G et que l'on arrive au point P_2 avec une vitesse nulle.

Naturellement

$$x_0 = \frac{R_1 + (d - R_2)}{2},$$

et il résulte d'une formule élémentaire, que la durée T du voyage entier de P_1 jusqu'à P_2 est

$$T = 2 \sqrt{\frac{d - R_1 - R_2}{G}}.$$

Il est possible, étant suffisamment loin des corps C_1 et C_2 , que les accélérations de gravitation de ces corps peuvent être négligeables à côté de l'accélération G , et de cette manière, la valeur absolue de l'accélération totale¹ observée ici par le voyageur

¹ Cette accélération totale signifie la somme de l'accélération cinématique et de celle de gravitation.

peut être approximativement égale à G , tandis qu'aux voisinages de C_1 et de C_2 la valeur absolue de l'accélération totale en question peut être bien supérieure à G , comme on le verra dans notre exemple numérique.

Maintenant, le problème principal de cet ouvrage est le suivant :

Problème 2 : Déterminer la durée du voyage de P_1 jusqu'à P_2 dans les mêmes conditions que celles du Problème 1, exceptée la condition relative à l'accélération, condition qui sera remplacée par la suivante, à savoir que ce n'est pas la valeur absolue de l'accélération cinématique, mais la valeur absolue de l'accélération totale observée par le voyageur qui est une constante donnée, et sa valeur est G^* . G^* signifie ici le maximum de l'accélération, dont l'action persistante est inoffensive pour l'organisme humain. Dans ces conditions nous cherchons la valeur de x_0 et la durée T^* du voyage entier de P_1 jusqu'à P_2 .

L'accélération totale observée par le voyageur est :

$$f\left(\frac{M_1}{x^2} - \frac{M_2}{(d-x)^2}\right) + \frac{d^2x}{dt^2},$$

(où f est la constante de gravitation : $f = 6,66 \cdot 10^{-8} \text{ g}^{-1} \text{ cm}^3 \text{ sec}^{-2}$), qui doit être égale à G^* resp. $-G^*$ dans les intervalles $R_1 < x < x_0$, resp. $x_0 < x < d - R_2$:

$$(1) \quad f\left(\frac{M_1}{x^2} - \frac{M_2}{(d-x)^2}\right) + \frac{d^2x}{dt^2} = G^* \quad (R_1 < x < x_0),$$

$$(2) \quad f\left(\frac{M_1}{x^2} - \frac{M_2}{(d-x)^2}\right) + \frac{d^2x}{dt^2} = -G^* \quad (x_0 < x < d - R_2).$$

En considérant le temps t comme la fonction de x , ces équations peuvent être résolues par rapport à $t = t(x)$, et on en tire :

$$(3) \quad t = \frac{1}{\sqrt{2}} \int_{R_1}^x \sqrt{\frac{x(d-x)}{G^*x^2(d-x) + f[M_1(d-x) + M_2x] + C_1x(d-x)}} dx \quad (R_1 < x < x_0),$$

$$(4) \quad T^* - t = \frac{1}{\sqrt{2}} \int_x^{d-R_2} \sqrt{\frac{x(d-x)}{-G^*x^2(d-x) + f[M_1(d-x) + M_2x] + c_2x(d-x)}} dx \quad (x_0 < x < d - R_2).$$

Les constantes d'intégration c_1 et c_2 peuvent être déterminées par les conditions

$$\left. \frac{dx}{dt} \right|_{x=R_1} = \left. \frac{dx}{dt} \right|_{x=d-R_2} = 0,$$

et il en résulte :

$$c_1 = -G^*R_1 - f\left(\frac{M_1}{R_1} + \frac{M_2}{d-R_1}\right),$$

$$c_2 = G^*(d - R_2) - f\left(\frac{M_1}{d-R_2} + \frac{M_2}{R_2}\right).$$

Il reste à déterminer la valeur de x_0 , qu'on peut faire en s'appuyant sur la condition suivante : la vitesse est une fonction continue même au point $x = x_0$, c'est-à-dire que

$$\left. \frac{dx}{dt} \right|_{x=x_0-0} = \left. \frac{dx}{dt} \right|_{x=x_0+0}.$$

On en tire

$$x_0 = \frac{d + R_1 - R_2}{2} + \frac{1}{2G^*} f\left(\frac{M_1}{R_1} - \frac{M_1}{d-R_2} + \frac{M_2}{d-R_1} - \frac{M_2}{R_2}\right).$$

En considérant qu'en général $R_1 \ll d$, $R_2 \ll d$, on peut écrire :

$$x_0 \sim \frac{d}{2} + \frac{1}{2G^*} f\left(\frac{M_1}{R_1} - \frac{M_2}{R_2}\right).$$

(Dans le cas où $M_1 = M_2$ et $R_1 = R_2$, il résulte $x_0 = \frac{d}{2}$).

Finalement la durée T^* du voyage entier est la suivante :

$$T^* = \frac{1}{\sqrt{2}} \int_{R_1}^{x_0} \sqrt{\frac{x(d-x)}{G^*x^2(d-x) + fM_1(d-x) + fM_2x + c_1x(d-x)}} dx + \\ + \frac{1}{\sqrt{2}} \int_{x_0}^{d-R_2} \sqrt{\frac{x(d-x)}{-G^*x^2(d-x) + fM_1(d-x) + fM_2x + c_2x(d-x)}} dx.$$

Pour avoir une notion sur les grandeurs de T et de T^* , citons l'exemple de la Terre et de Mars. Dans ce cas, on a :

$$\begin{aligned} R_1 &= 6380 \text{ km}, & M_1 &= 5970 \cdot 10^{21} \text{ kg}, \\ R_2 &= 3380 \text{ km}, & M_2 &= 639 \cdot 10^{21} \text{ kg}, \\ d &= 55 \cdot 10^6 \text{ km} \end{aligned}$$

(où d signifie la valeur approximative de la distance de la Terre à Mars au cas du périhélie²), et supposons que l'accélération cinématique G resp. l'accélération totale permise G^* sont égales à $\frac{3}{2}g$:

$$G = G^* = \frac{3}{2} \cdot 981 \text{ cm sec}^{-2}.$$

Dans ces conditions on obtient pour T et T^* :

$$T = 34,0 \text{ heures},$$

$$T^* < 35 \text{ heures}.$$

Ces résultats montrent que la durée T^* du voyage sous les conditions du Problème 2 ne dépasse pas de plus 3% la durée T du voyage sous les conditions du Problème 1.

Enfin, j'exprime mes remerciements à MM. les ingénieurs A. KRÁVITS et E. NAGY qui ont attiré mon attention sur le problème.

EGY MEGJEGYZÉS AZ ŰRHAJÓZÁSRÓL

ADLER GY.

Kivonat

A szerző avval a kérdéssel foglalkozik, hogy mekkora az a minimális időtartam, amely alatt egy égitestről egy másikra elérhetünk (nem tekintve a technikai akadályokat), ha az űrhajó utasára ható, az űrhajó kinematikai gyorsulásából és az égitestek gravitációs vonzásából adódó össz-gyorsulás értékének maximumát, amelyet az emberi szervezet huzamosabb ideig képes elviselni, előre megadjuk. A szerző megmutatja, hogy a Föld és a Mars esetében, midőn a Mars földközeli van, a két bolygót állócsillagnak tekintve, ez a minimális időtartam másfél napnál nem több.

² On peut vérifier aisément que dans ce cas les effets du Soleil et des autres planètes sont négligeables.

ЗАМЕЧАНИЕ О МЕЖПЛАНЕТНЫХ ПУТЕШЕСТВИЯХ

GY. ADLER

Резюме

Автор занимается следующим вопросом: каков минимальный промежуток времени, за который можно перенестись с одного небесного тела на другое (не принимая во внимание технические препятствия), если заранее задан максимум ускорения, получаемого за счёт кинематического ускорения воздушного корабля и от гравитационного притяжения небесных тел, которое человеческий организм может выдержать в течение длительного промежутка времени. Автор показывает, что в случае Земли и Марса, когда Марс находится на наименьшем расстоянии от Земли и эти планеты считаются неподвижными, этот максимальный промежуток времени не превосходит 1,5 дней.

A kiadásért felelős az Akadémiai Kiadó igazgatója

Műszaki felelős: Pataki Ferenc

Kézirat beérkezett: 1959. VII. 20. — Példányszám: 800. — Terjedelem: 8,25 (A/5) ív, 1 melléklet

49432/59 Akadémiai Nyomda, Budapest — Felelős vezető: Bernát György

СОДЕРЖАНИЕ

EGERVÁRY, J. †	107
ADLER, Gy.: Один новый тип граничных задач теплопроводности	109
ALMÁR, I.—BALÁZS, B.: Приближённый метод расчёта траектории космической ракеты проходящей вблизи Луны	129
BÉKÉSSY, A.—TÓTH, K.: Замечание к статье I. Almár и B. Balázs: «Приближённый метод расчёта траектории космической ракеты проходящей вблизи луны»	143
CSORDÁS, L.—MENOEZEL, Gy.: Замечания к статье P. Medgyessy: «Один механический функциональный синтезатор»	149
RÉNYI, A.: О последовательном и параллельном сцеплении автоклавов и о теории смещения	155
SZÜSZ, P.: О перестановке членов условно сходящихся рядов	167
VAS, É.: Об эффективности секвенциального критерия отношения вероятностей	171
ÁDÁM, A.: О двухполюсных электрических сетях, IV.	183
FÉNYES, T.: Применение операционного исчисления к решению обыкновенных разностно-дифференциальных уравнений	191

КРАТКИЕ СООБЩЕНИЯ

ADLER, Gy.: Замечание о межпланетных путешествиях	197
---	-----

INDEX

EGERVÁRY, J. †	107
ADLER, G.: Un type nouveau des problèmes aux limites de la conduction de la chaleur	109
ALMÁR, I.—BALÁZS, B.: Approximate method of plotting the orbit of a space rocket passing near the Moon	129
BÉKÉSSY, A.—TÓTH, K.: Remark on the paper of I. Almár and B. Balázs: „Approximate method of plotting the orbit of a space rocket passing near the Moon”	143
CSORDÁS, L.—MENCZEL, G.: Remarks on the paper of P. Medgyessy: „A mechanical functional synthesizer”	149
RÉNYI, A.: On serial and parallel coupling of autoclaves and on the theory of mixing	155
SZÜSZ, P.: Über die Umordnung bedingt konvergenter Reihen.....	167
VAS, É.: On the efficiency of the sequential probability ratio test.....	171
ÁDÁM, A.: Über zweipolige elektrische Netze, IV.	183
FÉNYES, T.: Anwendung der Operatorenrechnung von Mikusiński zur Lösung gewöhnlicher Differentialdifferenzgleichungen	191

SHORT COMMUNICATIONS

ADLER, G.: Une remarque sur la communication intercosmique.....	197
---	-----

307.801

A MAGYAR TUDOMÁNYOS AKADÉMIA

MATEMATIKAI
KUTATÓ INTÉZETÉNEK
KÖZLEMÉNYEI

IV. ÉVFOLYAM, 3—4. FÜZET

1959

★

ТРУДЫ
МАТЕМАТИЧЕСКОГО
ИНСТИТУТА

АКАДЕМИИ НАУК ВЕНГРИИ

ТОМ IV., ВЫПУСК 3—4.

1959

★

PUBLICATIONS
OF THE
MATHEMATICAL INSTITUTE

OF THE
HUNGARIAN ACADEMY OF SCIENCES

VOLUME IV., FASC. 3—4.

1959



1959

TARTALOMJEGYZÉK

FINNEY, D. J.: A kísérletezés rentabilitása.....	203
LINNIK, JU. V.: Öt előadás a valószínűségyszámítás és a számelmélet bizonyos problémáiról	225
GRÄTZER GY.—SCHMIDT E. T.: Egy asszociativitási tétel alternatív gyűrűkre.....	259
KREKÓ B.: Lineáris egyenletrendszerek megoldása szimplex módszerrel.....	265
RÓZSA P.—SÁRKÁNY GY.: Ellenáramú szétválasztó vegyipari alapl műveletek elméleti fokozatszámának meghatározásáról, II.	277
BOSZNAV Á.: Többszabadságfokú lengőrendszer gerjesztett lengései vizsgálatának egy módszere s néhány alkalmazása	299
CSÁKI E.: A Wilcoxon-statisztika két módosításáról	313
HASZPRA O.: Néhány megjegyzés nomogramok leolvasási hibáihoz	321
ADLER GY.: Szemcsés anyagok ellenáramú hűtése	327
ADLER GY.—FREUD G.: A Mikusiński-féle operátorszámítás alkalmazása egy parciális differenciálegyenlet megoldására	367
KOSIK P.—SALLAY M.—ZIMÁNYI M.: Hővezetési problémák összetett peremfeltételek esetén	377
RÉNYI A.: Összefüggő gráfokról, I.	385
RÉNYI A.: Szummációs eljárások valószínűségyszámítási interpretációja	389

A MAGYAR TUDOMÁNYOS AKADÉMIA

MATEMATIKAI
KUTATÓ INTÉZETÉNEK
KÖZLEMÉNYEI

IV. ÉVFOLYAM, 3—4. FÜZET

1959

★

ТРУДЫ
МАТЕМАТИЧЕСКОГО
ИНСТИТУТА

АКАДЕМИИ НАУК ВЕНГРИИ

ТОМ IV., ВЫПУСК 3—4.

1959

★

PUBLICATIONS
OF THE
MATHEMATICAL INSTITUTE

OF THE
HUNGARIAN ACADEMY OF SCIENCES

VOLUME IV., FASC. 3—4.

1959



1959

A MAGYAR TUDOMÁNYOS AKADÉMIA
MATEMATIKAI KUTATÓ INTÉZETÉNEK
KÖZLEMÉNYEI

„A MAGYAR TUDOMÁNYOS AKADÉMIA ALKALMAZOTT MATEMATIKAI INTÉZETÉNEK KÖZLEMÉNYEI”

című kiadványsorozat folytatása

SZERKESZTI: RÉNYI ALFRÉD

TECHNIKAI SZERKESZTŐK: RÉVÉSZ PÁL, BOGNÁR KATALIN

A SZERKESZTŐSÉG CÍME: MATEMATIKAI KUTATÓ INTÉZET, BUDAPEST, V. REÁLTANODA U. 13/15.

A MATEMATIKAI KUTATÓ INTÉZET KÖZLEMÉNYEI az Intézet tudományos eredményeit tartalmazó és egyéb matematikai, valamint a matematika gyakorlati alkalmazásával kapcsolatos dolgozatokat közölnek. A folyóirat minden évfolyama 4 füzetből áll és körülbelül 30 nyomdai terjedelmű. A dolgozatok vagy valamelyik világnyelven jelennek meg, magyar és még egy világnyelven írt részletes kivonattal, vagy pedig magyarul, két világnyelven írt részletes kivonattal. Közlésre szánt dolgozatokat kérjük 2 gépelt példányban kivonattal ellátva a szerkesztő címére küldeni (Budapest V. Reáltanoda u. 13/15.).

A folyóirat, amelynek első évfolyama 1956-ban jelent meg, folytatása a „Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei” című kiadványsorozatnak, amelynek összesen három kötete jelent meg: I. kötet (1952), II. kötet (1953), III. kötet (1954).

A MATEMATIKAI KUTATÓ INTÉZET KÖZLEMÉNYEINEK előfizetési ára évfolyamonként belföldi címre 50,— Ft, külföldi címre 70,— Ft. (Kötetenként 5 \$) Belföldi megrendelések az Akadémiai Kiadón keresztül adhatók fel (Budapest V. Alkotmány u. 21., Magyar Nemzeti Bank egyszámlaszám: 15.915.111—46), külföldi megrendelések a Posta Központi Hírlap Iroda útján eszközölhetők (Budapest V. József nádor tér 1., Magyar Nemzeti Bank egyszámlaszám: 61.257). A folyóirat egyes füzetei 15,— Ft-os árában az Akadémiai Könyvesboltban kaphatók (Budapest V. Váci u. 22.). Cserekapcsolatok felvétele érdekében kérjük az Intézet Könyvtárához fordulni (Budapest V. Reáltanoda u. 13/15.).

ТРУДЫ
МАТЕМАТИЧЕСКОГО ИНСТИТУТА

АКАДЕМИИ НАУК ВЕНГРИИ

Продолжение издания

„A MAGYAR TUDOMÁNYOS AKADÉMIA ALKALMAZOTT MATEMATIKAI INTÉZETÉNEK KÖZLEMÉNYEI”

РЕДАКТОР: ALFRÉD RÉNYI

ТЕХНИЧЕСКИЕ РЕДАКТОРЫ: ПÁL RÉVÉSZ, KATALIN BOGNÁR

АДРЕС: МАТЕМАТИЧЕСКИЙ ИНСТИТУТ, BUDAPEST, V. REÁLTANODA U. 13/15., ВЕНГРИЯ

В ТРУДАХ МАТЕМАТИЧЕСКОГО ИНСТИТУТА печатаются статьи, содержащие результаты научно-исследовательской работы Института, и другие математические работы, а также статьи, связанные с практическими приложениями математики. Каждый том журнала выходит в 4 выпусках и содержит приблизительно 30 печатных листов. Статьи публикуются либо на каком-нибудь мировом языке с подробным резюме на венгерском и каком-нибудь другом мировом языке, либо на венгерском языке с подробным резюме на двух мировых языках. Работы, предназначенные для опубликования, просим посылать в двух напечатанных на машинке экземплярах вместе с резюме в адрес редакции (Budapest, V. Reáltanoda u. 13/15.).

Журнал, первый том которого вышел в 1956 году, является продолжением издания «A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei» (Труды Института Прикладной Математики Академии Наук Венгрии) вышедшего в трёх томах: Том I. (1952), Том II. (1953), Том III. (1954).

Цена подписки на ТРУДЫ МАТЕМАТИЧЕСКОГО ИНСТИТУТА в Венгрии 50 форинтов, на заграничный адрес 70 форинтов (\$5 — за каждый том). Местные заказы принимает Издательство Академии Наук (Budapest V., Alkotmány u. 21., счёт Венгерского Национального Банка 15.915.111—46), заграничные заказы принимает Центральная Журнальная Контора Почты (Budapest V. József nádor tér 1., счёт Венгерского Национального Банка 61.257). По поводу отношения обмена просим обращаться к Библиотеке Института (Budapest V. Reáltanoda u. 13/15. Венгрия).

PUBLICATIONS
OF THE MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES

continuing the series

„A MAGYAR TUDOMÁNYOS AKADÉMIA ALKALMAZOTT MATEMATIKAI INTÉZETÉNEK KÖZLEMÉNYEI”

EDITOR: ALFRÉD RÉNYI

TECHNICAL EDITORS: PÁL RÉVÉSZ, KATALIN BOGNÁR

ADDRESS: MATHEMATICAL INSTITUTE, BUDAPEST, V., REÁLTANODA U. 13/15, HUNGARY

THE PUBLICATIONS OF THE MATHEMATICAL INSTITUTE are publishing papers containing the results of scientific work of the Institute and other mathematical papers and papers on the practical applications of mathematics. The journal is published quarterly, 4 fasciculi are forming a volume consisting of about 30 printed lists. The papers appear either in a world-language with abstract in Hungarian and in an other world-language, or in Hungarian with abstracts in two world-languages. Papers intended for publication in the journal should be sent to the editor in 2 typewritten copies, with an abstract.

The journal, the first-volume of which appeared in 1956, continues the series „A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei” (Publications of the Institut des Mathématiques Appliquées de l'Académie des Sciences de Hongrie) of which 3 volumes were published altogether: Volume I. (1952), Volume II. (1953), Volume III. (1954).

The price of a volume of the PUBLICATIONS OF THE MATHEMATICAL INSTITUTE is 50,— Ft to an address in Hungary and 70,— Ft (\$ 5.— The Volume) to abroad. Subscriptions can be made at the Académie Publishing House in Hungary (Budapest V. Alkotmány u. 21., single account number in the Hungarian National Bank 15.915.111—46) resp. at the Posta Központi Hírlap Iroda from abroad (Budapest V. József nádor tér 1., single-account number in the Hungarian National Bank 61.257). For establishing exchange relations please write to the Library of the Mathematical Institute (Budapest V. Reáltanoda u. 13/15., Hungary).

THE ECONOMIC EFFICIENCY OF EXPERIMENTATION¹

by

D. J. FINNEY F. R. S.²

1. Introduction

1.1. Until 35 years ago, the responsibility of the statistician in research was thought to consist only in the analysis of experimental data as a preliminary to their interpretation. The concept of experimental design was practically unknown. From about 1923, a series of remarkable papers by R. A. FISHER revolutionized this outlook. The wealth of detailed knowledge of particular designs that derived from FISHER's work is of great theoretical interest and of major importance to the satisfactory conduct of experiments; perhaps even more important for the effective use of experimental resources, however, is the recognition of two general principles that emerge from the theory and practice of design:

(i) The design of an experiment in great measure determines the form of statistical analysis appropriate to the results.

(ii) The success of an experiment in answering the questions that interest the experimenter, without excessive expenditure of time and resources, depends largely upon right choice of design.

1.2. In this paper, rather than discuss technical details of particular experimental designs, I have chosen to consider more general topics relating to the optimal specification of an experimental design for a particular purpose. A great variety of different questions arises, and I can do no more than outline their scope and give additional attention to one or two that I find especially interesting. Most of the questions cannot be formalized as well as those on the pure structure of designs, and answers usually require intimate knowledge of the particular field of application. Too often, therefore, they are settled solely by the experimenter, who may ask the statistician to suggest designs only within a fairly rigid specification of his own choosing. For the statistician to be sole arbiter would be equally unsatisfactory, for he might fail to see the full implications of his recommendations in the ultimate interpretation of the research. The great need is for close collaboration between experimenter and statistician. As a statistician, I naturally concentrate on the statistical point of view, but I know how easily this can lead me into impracticable proposals and how frequently compromise between a theoretical ideal and practical convenience is essential.

1.3. I am convinced that the statistician can and should make as valuable a contribution to the general planning of an experiment or of a whole programme

¹ Read at the Biometric Symposium 7. 9. 1959 (Budapest).

² University of Aberdeen.

of experiments as he does, perhaps with greater mathematical technicality, to details of design relating to the allocation of treatments to plots and to the statistical analysis of experimental records. In doing so, he will need to exercise considerable tact, for an experimenter who has not previously experienced benefits from consulting a statistician about all phases of his research may not unreasonably distrust, and even resent, criticism of his choice of treatments or of the general scope of his experiments by one who is not a specialist in that particular field of science.

2. Examples of Topics

2.1. Under the narrower heading of experimental design, the statistician is accustomed to discuss particular designs, such as Latin squares, balanced incomplete blocks, and factorial designs, as well as procedures of randomization, methods for the construction of solutions of combinatorial problems, confounding, and the analysis of variance and covariance and associated computational techniques. More recently, there have been important developments in sequential experimentation, and the methods of Box and others for the exploration of optimal conditions have perhaps done more to extend the understanding and use of experimental design than anything else published in the last 20 years. Much has to be decided before statistician or experimenter resolves on the use of one particular design. Too often, the important decisions are taken without conscious thought, and without realization of the consequences that they may have for the precision of results obtained in respect of important issues or even for the relevance of the results to the main subject of inquiry.

2.2. As examples of some of the topics that tend to be neglected in any purely formal presentation of experimental design, I list the following:

- (i) Use of controls.
- (ii) Number of treatments to be tested.
- (iii) Number of different factors in a factorial experiment.
- (iv) Number of levels of a factor.
- (v) Number of replicates of treatment combinations.
- (vi) Number of similar experiments on different sites, at different times, or on different sources of material.
- (vii) Definition of plots or other units to which experimental treatments are to be applied (shape and dimensions of field plots, species, age and other characteristics of experimental animals, use of repeated observations on same plot at intervals of time).
- (viii) Choice of suitable design.
- (ix) Specification of measurements to be made and techniques for making them.
- (x) Use of concomitant variates for improvement of precision.
- (xi) Relation between successive experiments on different phases of an experimental programme.

2.3. It would be easy to speak at length on each one of these topics, yet still only to have discussed aspects familiar by experience to one statistician. Although this might have some value as illustration, even though it did not amount to a comprehensive account, the time required would be excessive. I have therefore chosen a few special situations that will exemplify the statistician's line of argument rather than pretend to be comprehensive.

3. Internal and External Economy

3.1. Considerations that may broadly be described as economic are inseparable from any discussion of the relative merits of alternative experimental procedures. Although I have no intention of suggesting that scientific research is to be evaluated merely according to conventional economic principles, in any experimental science there will naturally be a desire to employ resources effectively in relation to the purpose of the research. The manner in which these resources are measured will vary from one situation to another, and in practice emphasis may have to be placed on *one* component of the resources because of the impossibility of measuring *all* on the same scale. For example, an able scientist is likely to have more ideas for investigations than he himself or those who work under him can use, and his chief concern may be economy of time, in order that he shall spend long enough on any one project to extract adequate and precise information but not so long that the rate of increase of knowledge is too small to justify continuation. In other circumstances the main factor to be considered may be the quantity of materials consumed, the area of land occupied, or the number of animals used. Whatever the limiting factor or factors may be, they can conveniently be regarded as measuring *costs*; I do not imply that cost must be assessed on a monetary scale, although such a scale is often the simplest way of expressing several components of cost commensurably.

3.2. Under the heading of the *internal economy* of experimentation may be grouped matters relating to the optimal planning of experiments for a specified total cost. Whether the research is a disinterested inquiry into natural phenomena or a problem of immediate practical importance to some industrial or biological technology, the aim of maximizing the returns from an experiment relative to its cost is reasonable, and in broad general terms much can be done by preliminary study of alternative plans. Difficulty arises when the aims of the research are not simply the estimation of one or two clearly defined quantities, because then the maximization cannot be seen in exact mathematical terms. "To estimate the bactericidal potency of *X*, a new drug, relative to an established standard, *S*" may not be easy, especially if a wide range of conditions is to be examined; nevertheless, the efficiency of alternative experimental plans can be discussed in some detail, the aim being maximum precision of estimation relative to the total number of test cultures or other experimental units to be expended. The efficiency of a project "To relate the chemical structures of a group of compounds to their bactericidal activities" could not be discussed in the same detail: the various conceivable findings are not simply related to the magnitude of experimentation, and the values of alternative outcomes cannot be expressed on a scale commensurable with that of cost. When emphasis is to be placed on the general advancement of knowledge rather than on the acquisition of a particular item of information, an investigator can seldom state in advance which aspects of his experiment are of primary interest; even more rarely can he list the conclusions to which he might be led and assign values to them.

3.3. In the past, statisticians have paid considerable attention to the internal economy of experimentation, but only in recent years has realization grown that it is proper, and indeed desirable, for them also to advise on what the total cost of a research project ought to be in relation to the value of results

expected from it. Whether or not justification for the amount of experimentation undertaken should be sought on these lines depends upon the nature of the project, and here the contrast between fundamental and technological research is intensified. FISHER ([16], Chapter 4) has stated essentially the same contrast in relation to significance tests and inductive reasoning on the one hand, decision and acceptance procedures on the other. He emphasized that the use of significance tests and fiducial distributions has the object of summarizing the evidence currently available from scientific experiments and of aiding the advancement of science through the accumulation of knowledge; the procedure is quite distinct from that appropriate to the taking of an irrevocable decision that concerns some matter of policy or recommending one of several alternative courses of action. "An important difference is that Decisions are final, while the state of opinion derived from a test of significance is provisional, and capable, not only of confirmation, but of revision. An acceptance procedure is devised for a whole series of cases. No particular thought is given to each case as it arises, nor is the tester's capacity for learning exercised. A test of significance on the other hand is intended to aid the process of learning by observational experience." It is fairly easy, logically though perhaps not practically, to devise a cost function for the loss incurred if an experiment leads to the recommendation to farmers of an inferior variety of wheat instead of the best; one cannot reasonably talk in quantitative terms about the loss consequent upon failing to discover a new fundamental principle of science, or upon a belief in a wrong theory.

3.4. None the less, many experiments and programmes of research are conducted for specific purposes that permit advance formulation of the value of the results. Although this is especially true in technological research, instances arise even in pure science where matters of technique must be explored as a preliminary to the main project; for example, alternative methods of sampling an insect population or of counting blood cells may need to be compared with a view to accepting one as standard in future fundamental research, the relative merits being assessed on the basis of accuracy per unit cost with costs possibly measured in time rather than in money. I shall today be concerned more with problems of scientific technology, and the manner in which the planning of experimentation should be influenced by economic considerations that go beyond merely the most efficient utilization of specified resources.

3.5. If the purpose of an experiment or groups of experiments is to estimate a parameter whose magnitude is of importance to some sector of the national economy, other things being equal any increase in the size or number of experiments will improve the precision of the estimate (Examples are mentioned in Section 6 below). In so far as an increase in precision enables policy based upon the estimate to be more correctly formulated, the loss to the economy consequent upon the policy being less beneficial than that based upon the unknown true value of the parameter will be reduced. Hence any increase in precision can be assigned a value on the scale of costs, which may be set against the cost of additional experimentation. Study of the optimal conditions, defined as maximizing the net gain, relates to the *external economy* of the system. The connexion with the theory of decision functions is close, though the approach is somewhat different.

3.6. I may summarize by saying that discussion of the internal economy of an investigation usually turns upon maximizing the precision of an estimate

(or the average precision of several estimates) for a specified cost, or minimizing the cost of attaining a specified precision. Other problems may occasionally be considered under this head. Discussion of the external economy, on the other hand, is usually concerned with evaluating the disadvantages to the whole economy in which the investigation is set that will result from imprecise estimates, and balancing the gains from reducing the imprecision against the cost of the extra experimentation necessary in order to achieve this reduction.

4. Number of Factors

4.1. Many experiments are planned with inadequate appreciation of what factorial design, confounding, and fractional replication can achieve. If an experiment intended for the study of two or three factors can have others incorporated, without seriously increasing the size of the experiment or seriously reducing the replication, additional information may be obtainable for a negligible additional labour or cost.

4.2. Consideration of experiments on factors at 2 levels enables the argument for additional factors to be forcefully expressed. An experiment on one factor would seldom be regarded as adequately replicated unless it had about 8 plots of each level. If two such experiments on the same subject-matter but with different factors were contemplated, 32 plots would be involved; by combining these into one 2^2 experiment, however, the same number of plots could be used to give information on the interaction of the factors as well as on the average of main affects. Randomized blocks of 4 might be used, instead of blocks of 2, but any loss in precision from the larger blocks would be offset by the greater replication (now sixteen-fold) on each factor. In this experiment, 21 out of 31 d. f. are used in the estimation of error variance and only 3 are assigned to treatment comparisons, a somewhat extravagant provision for error since the advantage of extra error degrees of freedom decreases markedly after the first 10. It would be an unimaginative experimenter who could not think of additional factors highly relevant to his investigation. Introduction of a third factor and arrangement as 2^3 in 4 blocks of 8 involves some loss in precision from the larger blocks, but now 7 d. f. are measuring treatment comparisons and 21 d. f. estimate error: still only one-quarter of the effort of the experimenter is spent directly on comparing treatments and three-quarters on assigning a measure of precision to the comparisons. Without change of block size, one or two more factors can be introduced, by confounding (3-factor and higher interactions only) and using remaining high order interactions for estimating error. This leads to either 10 d. f. for treatments (main effects and 2-factor interactions in a 2^4) and 18 d. f. for error or 15 d. f. for treatments (2^5 design) and 13 d. f. for error. Thus saturation of the experiment with factors, so making it a single replicate, still leaves almost half the effort expended on error estimation. Fractional replication provides a method of supersaturation; this is of restricted value in so small an experiment, but, if one 2-factor interaction can be sacrificed, a one-half replicate of 2^6 is possible with the same blocks of 8 plots, and now the allocation of d. f. is 20 for treatment main effects and 2-factor interactions, 8 for error.

4.3. FISHER [15] has shown how to allow for the inevitable imprecision of the sample estimate of variance in assessing the information available from an experiment. If an experiment has an error mean square s^2 with f degrees

of freedom, the precision of a contrast estimated as a difference between the means of two sets of r plots is

$$(1) \quad \frac{r(f+1)}{2s^2(f+3)}.$$

An index of information for the whole experiment, with main effects and 2-factor interactions all regarded as of equal value, might be defined as this quantity multiplied by the number of such treatment effects that can be estimated (FINNEY, [9]). Table 1 shows values of this index for the six experiments discussed above. Even allowing that the first requires only half as many plots as the others and that the first two have smaller blocks, the advantages of 2^5 and 2^6 designs appear substantial. Of course, this index should not be employed uncritically, as it makes no distinction in value between main effects and interactions nor does it take account of the occasional need for estimation of 3-factor or higher order interactions.

Table 1

Information on main effects and 2-factor interactions obtainable from 2^n experiments on 32 plots or less

No. of factors n	No. of plots	Block size	Replication r	Error d. f. f	Main effects and 2-factor interactions	Index of information
1	16	2	8	7	1	$3.2/s^2$
2	32	4	16	21	3	$22.0/s^2$
3	32	8	16	21	6	$44.0/s^2$
4	32	8	16	18*	10	$72.4/s^2$
5	32	8	16	13*	15	$105.0/s^2$
6	32	8	16	8*	20	$130.9/s^2$

* Here the error, wholly or in part, is estimated from interactions

4.4. The preceding paragraphs are not to be read as uncompromising advocacy of multifactorial single or fractional replicates for every purpose. The statistician must know the potentialities of factorial design and must put clearly before the experimenter a statement of the gains to be expected from variants of the original proposals; the experimenter has the last word. Limited experience and capabilities of the staff responsible for executing and recording experiments, or the restriction of interest to one or two well-defined questions, occasionally make very simple designs preferable to those that on paper appear more informative. Despite the merits of *successful* multifactorial fractional replicates, a simpler design that gives trustworthy results on a narrower front is preferable to a scheme too ambitious for the circumstances of the experiment. Often the inclusion of 5 factors in an experiment instead of the 2 initially suggested involves only slight increases in the labour of performing the experiment and of analysing the results; if the experiment has to be conducted under great stress, or with the aid of staff unaccustomed to anything other than the simplest experiments, the success of the whole may be jeopardized by the risk of mistakes being made, and 2^2 may be a wiser

choice than 2⁵. Although the statistician may be continually looking for opportunities of inserting as many factors as possible into experiments, he should not let the arguments that led to Table I be his only guide, but should temper his ambition with discretion.

5. Use of Concomitant Variates

5.1. In most experimental designs, block constraints are important, and one of the duties of a statistician is to assess the best system of blocking. Although much depends upon experience of the field of research, any logical classification of experimental units that is possible before the experiment begins and that is in the least likely to be associated with the final "yield" measurements deserves consideration as a basis for blocking. Often the number of choices is greater than the experiment can accommodate, even though recourse be had to Latin square and other designs that use two or more block systems simultaneously. The statistician learns to be suspicious of statements by experimenters that differences between days of experimentation or between alternative supplies of a reputedly standard material are negligible, but instead tests them whenever possible in records of former experiments; not infrequently, he finds that the efficiency of comparisons between treatments would be much improved if experiments were balanced in respect of them.

5.2. When a choice must be made between alternative classifications, other things being equal some preference should be given to qualitative characters as a basis for blocking, as quantitative characters can often be dealt with by covariance analysis. If an experiment is conducted in the belief that it does not matter which of four different persons is responsible for a certain operation (or which of four different supplies of a nominally standard drug is used), so that no attempt is made to balance the design over the four, and subsequently obvious differences appear, salvage can still be effected. One person, say P_1 , can be taken as standard; three dummy variates can be defined such that the first takes the value 1 for all yields from P_2 and zero for P_1, P_3, P_4 , a second is 1 for P_3 only and zero for the others, and a third is 1 for P_4 only. Covariance analysis on all the dummy variates simultaneously then adjusts all comparisons of treatment means to estimated equivalence in respect of P_1, P_2, P_3, P_4 ; the sum of squares removed from error by regression on the three variates reduces to the ordinary sum of squares between persons if the design is in fact perfectly balanced over the four. Not only is this process laborious, but estimation of the covariance adjustments inevitably involves some loss of information that would not have occurred if balance had been perfect. Hence the advantage lies with using such a classification as a basis for blocks if this be possible.

5.3. A quantitative character, such as initial weight of an animal or yield of a plot in any pre-experimental period, may be such that the relation of final yield to it (apart from treatment and block effects) is a regression of fairly simple type, perhaps linear or quadratic. Covariance analysis using this one concomitant variate (and its square if the regression is quadratic) then secures that comparisons of adjusted treatment means are made on terms of equivalence in respect of the concomitant. None the less, as KEMPTHORNE [19] has emphasized, strict validity of the procedure depends upon the exact truth of the regression model, an assumption that is avoided if groups of plots exactly or approximately equal in respect of the quantitative character are

given the status of block constraints. OUTHWAITE and RUTHERFORD [20] made an instructive re-examination of data first reported and analysed by FEDERER and SCHLOTTFELDT [7]. An experiment on plant growth was arranged in randomized blocks, each block being a single line of plots and successive blocks being parallel lines. Inspection of the yields suggested a trend along the line of the blocks, and OUTHWAITE and RUTHERFORD eliminated the trend by a multiple covariance analysis, using position of a plot in the block as one concomitant variate and successive integer powers of this value as other concomitants; by continuing to a polynomial of degree 7, they were able to eliminate all differences corresponding to the average positions of the 8 treatments within blocks, a procedure similar to the use of dummy variates mentioned in the preceding paragraph. Nevertheless, they found that the errors of estimation inherent in the covariance adjustments constituted about a 15% loss of information relative to a design in which Latin square constraints were used to balance treatments in respect of positions within the blocks. Here is an instance in which a quantitative (positional) character would have been better used as a basis for blocking than in a covariance analysis. Undoubtedly blocking is to be preferred when practicable, and covariance should be regarded as a device for use when the experiment cannot hold enough block constraints, when an unsuspected source of variation is seen to be associated with some character during or at the end of the experiment (the character itself having been assessed before treatments were applied or being for other reasons independent of the treatments), or for the adjustment of mistakes in design.³

5.4. One way of using a quantitative character as a basis for blocking is to stratify the plots according to the value of the character. For example, randomized blocks for t treatments might consist of the t plots with the highest values, the t next highest, and so on. Table 2 shows the design of an experiment on the comparison of 5 inocula of a plant virus in which this idea was extended to a Latin square (COX and COCHRAN, [2]); five plants formed the primary blocks (rows), but the relative sizes of five leaves on each plant were used as a secondary character for an orthogonal set of blocks (columns). A similar method of control can be used in the absence of any primary blocking system. Suppose that 5 treatments are to be tested on 25 experimental units, or plots, for each of which a concomitant variate has been measured before the experiment begins. The values of the concomitant can be listed in order of magnitude from highest to lowest, and the treatment of each plot then determined by writing successive rows of a 5×5 Latin square alongside this list; the square of Table 2 would give the series of treatments

A, E, D, C, B, C, D, A, B, E, B, ... C, E, A.

Analysis proceeds exactly as for a Latin square. The enthusiast for eliminating every shred of an effect of the independent variate can make covariance adjustments as well if he wishes! Papers by COX [1] and FELDT [8] are relevant to the whole of this discussion.

³ This last is of course something that should not occur, but methods are required for dealing with it. An example closely allied to the present discussion has been discussed by FINNEY and COPE [14].

Table 2

Allocation of five virus inoculations, using five leaves on each of five plants in a randomized Latin square design

Plant no.	Relative size of leaf				
	1	2	3	4	5
I	A	E	D	C	B
II	C	D	A	B	E
III	B	C	E	A	D
IV	E	A	B	D	C
V	D	B	C	E	A

5.5 A modified form of balance has sometimes been used, especially in animal experiments where perhaps initial weight is the quantitative character. Instead of random allocation to treatments, the sets of animals allotted to different treatments are so selected as to make all treatments have mean initial weights as nearly equal as possible. Quite apart from the subjective influences that enter as soon as true randomization is set aside, this procedure inflates variation *within* treatments at the expense of variation *between* treatments. Tests of significance and assessments of standard errors may be seriously biased unless the statistical analysis uses the covariance analysis that the design was presumably specially devised to avoid! The theory has been discussed elsewhere (FINNEY, [10]); this design appears to have no merits.

6. Number of Experiments

6.1 YATES [21] has discussed an elementary approach to the external economy of estimating the optimal rate of application of fertilizer to a particular crop; his aim is to sample a region in which the crop is grown by placing experiments at randomly distributed sites. If all experiments are to be of one simple standard pattern, the cost of a series of n can be expressed approximately as

$$(2) \quad A(n) = c + an,$$

where c represents fixed expenditure on the whole programme and a is the additional cost per experiment. Write ξ for the rate of fertilizer application per unit area that would maximize the net benefit, under the assumption that all farmers in the region adopt this rate. The yields in the n experiments can be averaged so as to estimate a response curve, from which X , an estimate of ξ , can be obtained. If this response curve, the regression of yield on amount of fertilizer per unit area, were quadratic, the reduction in net gain resulting from use of X instead of the unknown true value ξ would be proportional to $(\xi - X)^2$; for any other regression relation, the reduction in yield will still be of this form to the first order. Although $(\xi - X)^2$ is unknown, its expectation, $V(X)$, can be estimated from the experiments, and will be inversely proportional to n , say

$$(3) \quad V(X) = v/n.$$

The loss in yield per unit area from future use of X instead of ξ therefore has expectation $\lambda v/n$.

6.2 If the estimate X is used as a basis of fertilizer practice for a total area T , the expected total loss of crop from errors of estimation of ξ is

$$(4) \quad L(n) = \lambda vT/n.$$

If $A(n)$ has been measured on a scale of equivalent value of crop, the optimal n can be determined by minimizing $A(n) + L(n)$, and is

$$(5) \quad n^* = (\lambda vT/a)^{1/2}.$$

When instead of n^* experiments fn^* are performed (where f is either less or greater than 1), the expected net returns from application of the results of the research are reduced by

$$(6) \quad \frac{(f-1)^2 an^*}{f},$$

an amount which is small when f is near to 1. Consequently, if experimental resources have to be allocated between several different research programmes, each may be made somewhat smaller than its own optimal without serious loss to the whole economy. The theory can be modified in order to take account of the estimate X being used on crops for a number of future years.

6.3. The advantage to be gained by increasing n has from one point of view been underestimated. Knowledge of response curves for very many sites within the region might enable the region to be subdivided into districts within each of which the response-potential of the crop was more homogeneous. If this were followed by separate recommendations on rate of fertilizer application for the several districts, the total crop yield should be greater than if one average value were recommended throughout.

6.4. Greater logical and mathematical complexities appear when the need is to make a choice between two alternative procedures in the best possible way, instead of to estimate an optimal on a continuous scale. One problem of this kind, relevant to situations that occur in a variety of technological fields, has been studied by GRUNDY, HEALY and REES [17], GRUNDY REES and HEALY [18].

6.5. Suppose that the desirability of making a change in some standard process of technology is under investigation (e. g. the use of a new synthetic plastic for a particular type of electrical insulation, or the incorporation of a certain hormone into the diet of young pigs). A unit experiment has been performed, from which the increase in the amount or quality of production by the new process relative to that by the old is estimated to be x with variance σ^2 ; x , of course, may be positive or negative and the variance will be supposed based on enough degrees of freedom for it to be taken as the population value. Then x is an estimate of a population parameter, θ representing the expected improvement in production attributable to the new process. A decision on whether or not to adopt the new process ought to depend upon whether or not $W\theta$, a measure of the advantage gained from the change (assumed proportional to θ), exceeds a , the costs inherent in making the change (costs of new equipment, loss of production during a period of change, additional labour requirements until a new routine is working satisfactorily, and so on). The investigator has two reasonable alternatives open to him:

- (i) Calculate $Wx - a$, and recommend the adoption or rejection of the new process according as this is positive or negative;
- (ii) Recognize that evidence is inadequate and conduct a further n units of experimentation so as to obtain a second estimate, y , of θ with variance σ^2/n . Then calculate

$$\frac{W(x + ny) - a}{1 + n},$$

and recommend adoption or rejection according as this is positive or negative.

6.6. If the cost of the additional experiments is proportional to their number, being a say, the expected gain from rule (ii) is a function of θ , x , and n :

$$(7) \quad Q(\theta, x, n) = (W\theta - a)P - an,$$

where P is the probability that the new process is adopted. Of course, Q may be negative, for example if n is taken excessively large. Suppose now that x and y are normally and independently distributed about their mean θ . The condition that rule (ii) leads to adoption of the new process may be written

$$y > -\frac{x}{n} + \frac{a(n+1)}{Wn};$$

the probability of this is

$$(8) \quad P = \Phi \{[nW\theta + Wx - a(n+1)]/\sigma Wn^{\frac{1}{2}}\},$$

where

$$(9) \quad \varphi(z) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}z^2}$$

and

$$(10) \quad \Phi(z) = \int_{-\infty}^z \varphi(t) dt.$$

Moreover, (8) can also be regarded as applicable under the conditions of rule (i), for as $n \rightarrow 0$, $P \rightarrow 1$ or 0 according as $(Wx - a)$ is positive or negative.

6.7. The investigator must decide what value of $n (\geq 0)$ he will use. If his decision is to be in some sense optimal from the economic point of view presumably it should be based only upon the behaviour of the function in (7). Now in $Q(\theta, x, n)$, x is known from the first experiment, but θ is unknown, so that the obvious course of maximizing Q is not available; n must be chosen solely as a function of x and of the parameters of costs and variability. One principle that has found favour in problems of this kind is that of the *minimax*, which involves choosing n as a function of x in such a way as to minimize the maximum loss that can occur through the value of θ being unfavourable to the course of action that is eventually adopted. This loss is measured relative to the value of an immediate correct decision, $(W\theta - a)$ or 0 according as $(W\theta - a)$ is positive or negative; the loss is the difference between this amount and $Q(\theta, x, n)$, and the minimax value of n is that which minimizes the maxi-

mum of the loss function with respect to θ . Minimax estimation is known to have certain optimal properties, but GRUNDY and his colleagues remark that they "are not aware of any necessity for preferring the minimax solution to all alternatives in practical problems". In this problem, the minimax method is mathematically intractable; they propose instead to choose n so as to maximize a value of \bar{Q} averaged with respect to θ , employing for this purpose the fiducial distribution of θ on the evidence of the first experiment, that is to say a normal distribution of mean x and variance σ^2 . The average is therefore

$$(11) \quad \begin{aligned} \bar{Q}(x, n) &= \int_{-\infty}^{\infty} Q(\theta, x, n) \varphi\left(\frac{\theta - x}{\sigma}\right) \sigma^{-1} d\theta \\ &= \int_{-\infty}^{\infty} (Wz\sigma + Wx - \alpha) \Phi\left[zn^{\frac{1}{2}} + \frac{(Wx - \alpha)(n + 1)}{\sigma Wn^{\frac{1}{2}}}\right] \varphi(z) dz - \alpha n \end{aligned}$$

by substitution from (7) and (8) and the transformation $\theta = z\sigma + x$. Some manipulation of standard integrals then leads to

$$(12) \quad \bar{Q}(x, n) = (Wx - \alpha) \Phi(u) + \sigma W \left(\frac{n}{n + 1}\right)^{\frac{1}{2}} \varphi(u) - \alpha n,$$

where

$$(13) \quad u = \frac{(Wx - \alpha)}{\sigma W} \cdot \left(\frac{n + 1}{n}\right)^{\frac{1}{2}}.$$

Note that, as $n \rightarrow 0$, \bar{Q} tends to $(Wx - \alpha)$ or 0 according as $(Wx - \alpha)$ is positive or negative, and so gives correctly $\bar{Q}(x, 0)$.

6.8. It is convenient to regard n as measured on a continuous scale, so making it the ratio of the amount of experimentation recommended for the second stage to that already undertaken in the determination of x . Differentiation of $\bar{Q}(x, n)$ with respect to n shows that, as n increases from zero, \bar{Q} first decreases and may thereafter either continue to decrease or attain first a minimum, then a maximum, and subsequently decrease steadily. Practical importance attaches to the absolute maximum of \bar{Q} for $n \geq 0$; the recommendation will be to take an immediate decision ($n = 0$) if $\bar{Q}(x, 0)$ is not exceeded by $\bar{Q}(x, n)$ for any $n > 0$, but to use the value of n corresponding to the maximum of \bar{Q} if this exists and exceeds $\bar{Q}(x, 0)$. GRUNDY *et al* have studied the conditions for an absolute maximum of \bar{Q} and have constructed a nomogram for determining the optimal n .

6.9. Enough has been said to illustrate the character of this theory. The authors have developed the mathematics in greater detail, have given tables to show how the recommendations operate for particular values of x and the various parameters, and have compared the performance of their method with that of several alternatives. As might be expected, the function $\bar{Q}(x, n)$ is flat in the neighbourhood of its maximum, so that little harm comes

of taking n at some distance from its maximizing value. Moreover (except when $Wx - a = 0$), the behaviour of \bar{Q} in the neighbourhood of $n = 0$ is such as to ensure that either $n = 0$ is recommended or the recommended additional amount of experimentation is fairly large; clearly a small amount is of little use as it can neither confirm nor controvert the evidence of the quantity $(Wx - a)$ effectively and is therefore almost sure to be economically disadvantageous.

7. Relation between Successive Experimental Phases

7.1. A major research project will seldom involve only one experiment or one group of contemporaneous experiments. Different aspects of the research are likely to involve different schemes of experimentation, related by the fact of being part of one project. Important questions arise in connexion with the allocation of effort between these phases; one class of problems of considerable statistical interest is that in which the "treatments" on which experiments are to be conducted in one phase depend upon the outcome of an earlier phase. I shall discuss two examples of the efficient planning of selection experiments, in which an initially large number of treatments are to be subjected to experiment, as a result of which the best performers will be selected for a second stage of experimentation.

7.2. I first consider a basic problem of plant breeding for yield improvement, although the mathematical model that I adopt is necessarily somewhat crude; I have discussed elsewhere (FINNEY, [11], [12]) numerous practical points that need to be taken into account. By crossing established varieties, plant breeders can produce large numbers of new seedlings of a crop species; the great majority of these will prove to be useless, but some may deserve perpetuation as the foundations of new varieties. Suppose that each year a cohort of N potentially new varieties of a crop is ready to begin its programme of yield testing, that testing is to continue over k successive years, and that at the end of this time a proportion π of the varieties is to be passed forward as "successes". The number of varieties of the cohort retained under test will be reduced from year to year, so that in year r all survivors from year $(r-1)$ will be grown in a field trial and the fraction P_r of these showing the best yields (without any reference to tests of significance) will be retained for a further year. Account must be taken of the possibility that N may be so large as to make the testing of all in the first year inexpedient, and instead P_0N might be randomly selected from the cohort, the remaining fraction $(1-P_0)$ being discarded without test. Clearly the P_r are subject to the constraint

$$(14) \quad P_0 P_1 P_2 \dots P_{k-1} P_k = \pi.$$

7.3. In any one year, survivors of k different cohorts will be under test, each at a different stage of testing. Under stable conditions the experimenter, will have a total area A available for field trials in each calendar year, the site perhaps changing from year to year although its area is constant. This he must subdivide so as to allot an area A_r to the $P_0 P_1 \dots P_r N$ survivors of the cohort now in its r th year of testing ($r = 1, 2, \dots, k$), where

$$(15) \quad A_1 + A_2 + \dots + A_{k-1} + A_k = A.$$

The problem of optimal planning is to determine values of P_0 , the P_r , and the A_r , subject to (14), (15), so as to maximize the expected yields of the πN varieties finally selected.

7.4. The N varieties may be regarded as a sample from an infinite population in which the distribution of expected yields is of specified form, and only a normal distribution will be considered here; CURNOW [3], [4] has obtained some results pertaining to other forms of distribution. The field trial at stage r will estimate the expected yields with an error that may reasonably be assumed normal; moreover, this error will decrease as the number of varieties in stage r is decreased or as A_r is increased, on account of changes in plot size and replication that are made possible. As an approximation, ε^2 , the error variance will be assumed to be expressible as

$$(16) \quad \varepsilon^2 = \frac{\gamma P_0 P_1 \dots P_{r-1} A \sigma^2}{A_r},$$

where γ is a constant and σ^2 is the variance of the distribution of expected yields.

7.5. Mathematical analysis of this model for one-stage selection is then quite simple. The mean yield of the πN varieties selected can be shown to have an expectation that exceeds the general mean by an amount

$$(17) \quad G = \sigma Z_1 / P_1 (1 + \gamma P_0)^{\frac{1}{2}},$$

where Z_1 is the ordinate to the standardized normal frequency function (mean 0, variance 1) corresponding to a single-tail probability P_1 . Under condition (14) this is maximized by taking P_1 as the solution of

$$Z \ 2P + \pi \gamma = 2TP (P + \pi \gamma),$$

where T is the abscissa (unit normal deviate) and Z the ordinate corresponding to P . Hence P_1 is a function of $\pi \gamma$ alone, and $P_0 = \pi / P_1$. Equation (18) can lead to a value of P_1 smaller than π , which is an indication that G , the gain in mean yield, would have been greater if N had been greater but that the best procedure now is to take $P_1 = \pi$, $P_0 = 1$. More commonly, P_1 will be greater than π , and an initial random discard of $N(1 - P_0)$ varieties will be advantageous. The gain can be quite considerable; for example, if $\gamma = 5$ and $\pi = 0.01$, reasonable practical values, the optimal P_1 is 0.063, and initial reduction of the N varieties to 0.16 N will increase G by 35% as compared with using $P_1 = \pi$ on all the original varieties; if π had been 0.1, the gain from using the optimal procedure, a random discard to reduce N to 0.56 N followed by $P_1 = 0.18$, would have been only 5% of the value of G for $P = 0.1$.

7.6. As soon as more than one stage has to be considered, the mathematics increase in complexity, and even for two-stage selection little has yet been achieved beyond numerical study of one or two particular cases. However, these are quite revealing. As might be expected, an initial discard is much less important, since its function of reducing the total number of varieties to a number that can be tested with reasonable precision on the available land is in part performed by the first stage of selection. If π is very small, a value of P_0 different from unity may be advantageous, but a gain of 10% or more relative to the best that can be achieved with $P_0 = 1$ is exceptional. Only the

case of $P_0 = 1$, therefore, has been studied in any detail. For any specified values of γ and π , a selection programme is then completely defined by choice of P_1 and A_1 , since $P_2 = \pi/P_1$, $A_2/A = 1 - A_1/A$. If two rectangular axes are taken as scales of P_1 (better, $\log P_1$) and A_1/A , the quantity G can be evaluated for many different pairs of points in the diagram; G is again defined to be the excess of the expected mean yield of the πN selected varieties over the general mean of all varieties, but its symbolic expression is much more cumbersome than (17) and is laborious to compute. Contours of equal G can be inserted on the diagram, and these will surround a single maximum at the optimal combination of P_1 and A_1 . A full study of $\gamma = 1$, $\pi = 0.01$, supported by a lesser set of computations for $\gamma = 10$, $\pi = 0.01$ and argument from the run of the results suggests a simple broad generalization: under a wide range of values of γ and π corresponding roughly to practical conditions, the maximization of G is achieved approximately by

$$(19) \quad P_1 = P_2 = \pi^{\frac{1}{3}}, \quad A_1 = A_2 = \frac{1}{2} A.$$

These certainly do not give the true maximum, but the surface relating G to P_1 and A_1 is very flat near its maximum and the loss from the approximation is negligible. When π exceeds 0.1, this two-stage selection is scarcely any better than one-stage; as π decreases the advantage of the extra stage becomes greater, although it appears that the gain relative to the optimal one-stage procedure is always much smaller than that for optimal one-stage relative to one-stage with $P_0 = 1$. CURNOW [5] has made some calculations relating to three-stage selection, which support the view that

$$(20) \quad \left. \begin{aligned} P_0 &= 1, \\ P_1 &= P_2 = P_3 = \pi^{1/3}, \\ A_1 &= A_2 = A_3 = A/3, \end{aligned} \right\}$$

approximate to the optimal conditions and also indicate that the further gain relative to two-stage is small. The generalization of this simple rule to k stages is obvious; although the deviation of the approximate values of the P , and A , from the true maximizing values may increase with increasing k , the difference between the mean expected yields for those varieties selected according to the rule and those selected according to a true optimal procedure is almost certainly always small.

7.7. CURNOW [3] and I [13] have made some progress with study of the external economy associated with this problem of selection of crop varieties. For one-stage selection, the mathematics are relatively simple and curves can be constructed from which the optimal policy can be derived; for two-stage the analysis is more complex. Despite the weaknesses of the model, this work does give some guidance on the amount of effort that can most advantageously be spent on varietal selection.

7.8. Instead of discussing this, I shall conclude with an account of another selection problem, that of the "screening" of chemical compounds for possible therapeutic activity. During the manufacture of pharmaceutical and other chemical products, large numbers of different compounds may be made; any one may be of value in the treatment of a particular disease, but the proportion of active compounds will be exceedingly small and only empirical

test can determine whether or not a compound is active. A cure for cancer may already exist as a compound synthesized for an entirely different research or industrial purpose and since put aside as of no further immediate interest, yet untried in a field where it will prove immensely valuable.

7.9. Here again, there is need for preliminary tests with low replication, on the basis of which compounds whose therapeutic activity is very slight can be rejected, followed by tests with higher replication as a second stage; again the possibility of a sequence of several stages can be considered. In this work, however, a better approximation to reality than can be given by a continuous distribution of expected yields may be provided by a discrete distribution with an exceedingly small proportion of active compounds (assumed of equal value) and a complementary large proportion of compounds with zero activity.

7.10. DAVIES [6] has made a beginning in the development of statistical theory appropriate to this situation. Necessarily he restricted himself to the supposition that the compounds tested can be regarded as randomly selected from a large population. Although non-statistical considerations must be taken into account, a plan of screening derived from a theory based on randomness is likely to be useful as a first approximation to an optimal procedure. DAVIES remarked that "When an active compound has been found, it is usual to make a number of compounds of similar value in the hope of improving the activity. This is what we call 'following a lead'. When a lead is being followed the test is no longer a random one, and the considerations on which to base the design of the test are then different".

7.11. For many purposes, the original distribution of activity may be taken as having only two values, 1 and 0, with probabilities δ and $(1 - \delta)$ where δ is usually small (possibly of the order of 0.01, 0.001, or even smaller). Any screening procedure will involve testing compounds, perhaps by giving each to one or more animals, measuring the activity, and selecting those for which the mean activity so measured exceeds a specified value; the measurement will be subject to experimental error, and so will not be restricted to the values 1 and 0. Inevitably some of those selected will be *false positives*, that is to say compounds with true value 0 that pass the test by the chances of experimental error, and equally inevitably some true positives will fail to be selected. DAVIES suggested as a suitable criterion for an optimal scheme that, when the total number of compounds is reduced by selection to a fixed proportion, π , of itself, the proportion of truly active compounds should be maximized (or the proportion of false positives minimized) relative to the total effort expended. For this initially binomial distribution of true activity, the criterion coincides with that of maximizing the true mean activity; for a more general distribution, a criterion based upon the proportion of compounds that exceed a certain level of true activity appears to be preferable to one based upon the mean.

7.12. Suppose that the rules of selection adopted are such that compounds with true activities 0, 1 have probabilities θ , θ' respectively of being selected. In order to conform to the requirement that the intensity of selection shall be π , these must satisfy

$$(21) \quad \pi = \theta (1 - \delta) + \theta' \delta.$$

The proportion of active compounds amongst all compounds selected is then

$$(22) \quad \psi = \theta' \delta / \pi,$$

and ψ is also the mean activity of those selected. The testing will involve giving doses of the compounds to experimental units, these perhaps being members of one of the usual species of laboratory animal. If the procedure has more than one stage, the number of animals used may vary from one compound to another, and in particular may differ in average between active and inactive compounds. If the expectations of the numbers of animals used for compounds of activity 0, 1 are m , m' respectively,

$$(23) \quad \bar{m} = m(1 - \delta) + m' \delta$$

is the average number of animals per compound. The expenditure in testing one compound can perhaps be satisfactorily represented as

$$(24) \quad A = a\bar{m} + c,$$

where a is the cost per animal used and c the cost of the compound itself. The aim will therefore be to maximize ψ subject to π and A being fixed, and with δ unknown but presumed very small. These expressions are easily generalized to integrals if the original distribution of activity is continuous.

7.13. One possibility would be to adopt a k -stage testing programme analogous to that discussed earlier in this Section. If $k = 2$, computations can be expeditiously performed with the help of existing tables of the normal distribution. Suppose that the first stage assigns r_1 animals to each compound; the observed effect of any compound has a mean x whose variance is σ^2/r_1 , where σ^2 is the variance per animal. All compounds are rejected for which $x < X$, and this amounts to selection of a proportion P_1 , so reducing an initial N compounds to P_1N . At the second stage, r_2 animals are assigned to each remaining compound and the mean effect, y , has variance σ^2/r_2 . All compounds are rejected for which the mean over both stages is less than Y , that is to say

$$r_1x + r_2y < (r_1 + r_2) Y,$$

$$(25) \quad y < [(r_1 + r_2) Y - r_1x]/r_2,$$

and this amounts to selection at the second stage of a proportion P_2 , where

$$(26) \quad P_1P_2 = \pi.$$

Here X and Y are quantities still to be determined, subject to (26).

7.14. If normality of distribution can be assumed for x and y , the probability that a compound is selected at the first stage is

$$(27) \quad \theta_1, \theta'_1 = \left(\frac{r_1}{2\pi\sigma^2} \right)^{\frac{1}{2}} \int_{\frac{(X-\mu)\sqrt{r_1}}{\sigma}}^{\infty} e^{-\frac{1}{2}t^2} dt,$$

where μ takes the values 0, 1 for the two values of the true activity. The probability that a compound that goes forward from the first stage is selected at the second stage is the probability that y should exceed $[(r_1 + r_2) Y - r_1x]/r_2$ given that x exceeds X , and this can be shown to be

$$(28) \quad \theta_2, \theta'_2 = H \left[\frac{(X - \mu)\sqrt{r_1}}{\sigma}, \frac{(Y - \mu)\sqrt{(r_1 + r_2)}}{\sigma}, \sqrt{\frac{r_1}{r_1 + r_2}} \right],$$

where $H[\xi, \eta, \rho]$ symbolizes the standardized bivariate normal integral:

$$(29) \quad H[\xi, \eta, \rho] = \frac{1}{2\pi(1-\rho^2)} \int_{\xi}^{\infty} \int_{\eta}^{\infty} \exp \cdot [-(u^2 - 2\rho uv + v^2)/2(1-\rho^2)] \, du \, dv$$

and again $\mu = 0, 1$ are to be substituted. Then in the notation of the earlier part of this Section,

$$(30) \quad \begin{aligned} \theta &= \theta_1 \theta_2, \\ \theta' &= \theta'_1 \theta'_2, \end{aligned}$$

and

$$(31) \quad \begin{aligned} m &= r_1 + \theta_1 r_2, \\ m' &= r_1 + \theta'_1 r_2. \end{aligned}$$

7.15. The investigation of optimal conditions is much simplified by making use of the fact that δ is very small, so that, from (21), θ is almost equal to π . Particular conditions can be examined by choosing values for θ_1 and r_1 (an integer). If σ is known with sufficient accuracy from previous experience, tables of the normal integral can be used to give \bar{X} from (27), and also to give θ'_1 by use of the other value of μ . Hence approximately θ_2 is found as

$$\theta_2 = \pi/\theta_1.$$

Corresponding to any trial value of r_2 , Y can be determined from (28) by use of tables of the bivariate normal integral, and (28) also gives θ'_2 . From (30) θ' is obtained. Because δ is small, the mean number of animals per compound will scarcely differ from m , or $(r_1 + \theta_1 r_2)$. Finally

$$(32) \quad \frac{\theta'}{A} = \frac{\theta'}{a(r_1 + \theta_1 r_2) + c}$$

is evaluated, which by (21), (22) is proportional to the frequency of true positives per unit cost amongst the selected compounds. Repetition of the cycle of calculation for alternative values of r_1, r_2, θ_1 , enables a search for the maximum of (32) to be conducted.

7.16. Enough has been said to indicate Davies's approach to this problem, although the account given is only a simplified version of his. He gave a numerical illustration of the calculations in connexion with the screening of compounds that might increase survival time amongst mice infected with tuberculosis. He compared both the theory and the particular practical application of the two-stage test with a full sequential procedure, and found the latter to be considerably more efficient in respect of frequency of positives selected per unit cost. However, a sequential method is sometimes inconvenient in practice, and he suggested that a three stage selection scheme might be a satisfactory compromise; as far as is known, no research has been done on this.

(Received June 2, 1959.)

REFERENCES

- [1] COX, D. R.: "The use of a concomitant variable in selecting an experimental design". *Biometrika* **44** (1957) 150—158.
- [2] COX, G. M.—COCHRAN, W. G.: "Designs of greenhouse experiments for statistical analysis". *Soil Science* **62** (1946) 87—98.
- [3] CURNOW, R. N.: *The Statistical Theory of Selection* Ph. D. Thesis — (1959) University of Aberdeen.
- [4] CURNOW, R. N.: "The consequences of errors of measurement for selection from certain non-normal distributions". *Bulletin de l'Institut International de Statistique* **37** (1960a).
- [5] CURNOW, R. N.: (1960b) (paper in preparation).
- [6] DAVIES, O. L.: "The design of screening tests in the pharmaceutical industry". *Bulletin de l'Institut International de Statistique* **36** (1958) 226—241.
- [7] FEDERER, W. T.—SCHLOTTFELDT, C. S.: "The use of covariance to control gradients in experiments". *Biometrics* **10** (1954) 282—290.
- [8] FELDT, L. S.: "A comparison of the precision of three experimental designs employing a concomitant variable". *Psychometrika* **23** (1958) 335—353.
- [9] FINNEY, D. J.: "The statistician and the planning of field experiments". *Journal of the Royal Statistical Society, A* **119** (1956) 1—27.
- [10] FINNEY, D. J.: "Stratification, balance, and covariance". *Biometrics* **13** (1957) 373—386.
- [11] FINNEY, D. J.: "Plant selection for yield improvement". *Euphytica* **7** (1958) 83—106.
- [12] FINNEY, D. J.: "Statistical problems of plant selection". *Bulletin de l'Institut International de Statistique* **36** (3) (1958) 242—268.
- [13] FINNEY, D. J.: "A simple example of the external economy of varietal selection". *Bulletin de l'Institut International de Statistique* **37** (1960).
- [14] FINNEY, D. J.—COPE, F. W.: "The statistical analysis of a complex experiment involving unintentional constraints". *Biometrics* **12** (1956) 345—368.
- [15] FISHER, R. A.: *The Design of Experiments* (6th edition). London (1951). Oliver & Boyd.
- [16] FISHER, R. A.: *Statistical Methods and Scientific Inference*. London (1956). Oliver & Boyd.
- [17] GRUNDY, P. M.—HEALY, M. J. R.—REES, D. H.: "Economic choice of the amount of experimentation". *Journal of the Royal Statistical Society B* **18** (1956) 32—55.
- [18] GRUNDY, P. M.—REES, D. H.—HEALY, M. J. R.: "Decision between two alternatives — How many experiments?" *Biometrics* **10** (1954) 317—323.
- [19] KEMPTHORNE, O.: *The Design and Analysis of Experiments*. New York (1952). Wiley & Sons.
- [20] OUTHWAITE, A. D.—RUTHERFORD, A. A.: "Covariance analysis as an alternative to stratification in the control of gradients". *Biometrics* **11** (1955) 431—440.
- [21] YATES, F.: "Principles governing the amount of experimentation in developmental work". *Nature* **120** (1952) 138.

A KÍSÉRLETEZÉS RENTABILITÁSA

D. J. FINNEY

Kivonat

A statisztikus feladata manapság nemcsak abban áll, hogy a kísérleti eredményeket analizálja, hanem hogy a kísérletek tervezésében is segítséget nyújtson. Szükséges annak felismerése, hogy a szűkebb értelemben vett kísérleti tervezés (design) csak egyik szempontja a tágabb értelemben vett kísérleti tervezésnek és nem szabad, hogy a statisztikus részvétele csak a tervezés leszűkített kombinatorikus problémáira szorítkozzék. Ahhoz, hogy a kísérletező erőfeszítései a lehető leghatásosabbak legyenek, kell, hogy szoros együttműködés legyen a statisztikus és kísérletező tudós között.

Az előadásban szó volt a kísérletezésnek mind a belső, mind a külső ökonómiájáról, tehát arról a kérdésről, hogyan kell a kísérletet úgy tervezni, hogy a rendelkezésre álló időt, anyagot, munkát a legnagyobb hatásfokkal használjuk fel, továbbá arról, hogy mekkora lehetőségeknek kell rendelkezésre állni ahhoz, hogy a kísérlet a legjobban szolgálhassa a kísérlet célját. Az alapvető kutatásoknál gyakran csak a belső ökonómiát érdemes vizsgálni, de sok technológiai kísérletnél a külső tényezőket is tekintetbe kell venni. A második fejezetben felsoroltunk egy-két olyan különleges kérdést, melyeknek megvitatása szükséges.

A 4—7. fejezetek a tervezés egyes speciális problémáinak elméletéről és gyakorlatáról szólnak. Először a vizsgált faktorok optimális számának meghatározásáról esett szó. A szerző véleménye szerint minőségileg és mennyiségileg is az a legelőnyösebb, ha olyan sok faktort vonunk egyszerre be a vizsgálatba, amennyit csak megenged a kísérlet mérete. A második kérdéscsoport a parcellákra és más kísérleti egységekre vonatkozó concomitans információ lehetőségének problémája. Ennek viszonylagos előnye akkor mutatkozik meg, amikor a homogén blokkok képzéséhez használjuk fel, vagy amikor az eredményeket a concomitanssal végzett kovariancia analízis segítségével korrigáljuk.

A harmadik probléma arról szól, hogy miként lehet meghatározni egy adott tényező tisztázására irányuló kísérletezés szükséges mennyiségét. Az egyik fajtája ennek az, amikor meg kell határozni e hasonló kísérletek optimális számát olyan esetben, amikor az a célunk, hogy praktikus használatra megállapítsuk valamely faktor legjobb szintjét (pl. a vetés alatt álló területegységre jutó műtrágya mennyiségét). Ez matematikailag sokkal egyszerűbb, mint amikor olyan kísérletek optimális számát kell meghatároznunk, melyek alapján gyakorlati döntést kell hoznunk abban, hogy két alternatív és élesen különböző eljárás közül melyiket válasszuk (pl. két különböző diéta). Az előadás érintette a második fajta problémát is olyan esetekre, mikor van bizonyos előzetes információnk és a kísérletezőnek az eddigi adatok alapján, vagy azonnal kell döntenie, vagy meg kell határoznia a döntéshez szükséges további kísérletek mennyiségét.

Végül a szerző két példával illusztrálja az egymást követő lépcsőkből álló kísérletezések optimális tervezésének kérdését. Az egyik a természetben növények varietásainak szelekciójával foglalkozik, amikor is az egymást követő idényben végeznek termőföldi kísérleteket és minden idényben a legjobbaknak egy hányadát választják ki a következő idényben történő folytatálagos kísérlethez. Ha a varietások kezdeti száma és az utolsó idény végén is megtartottak száma továbbá a teljes kísérletezési terület adott, a szelekciós frakciók arányát és az egyes stádiumokhoz szükséges földterületek arányát kell megállapítani. Láthatóan az „egyenlő frakciók minden évben és az egész területnek az egyenlő elosztása” adja azt a sémát, mely általában közel áll az optimálishoz. A második példa arról szól, hogy igen nagy számú kémiai preparátumot szűrünk azért, hogy megkeressük azokat, melyeknek therapiás értékük van. A kezdeti populáció valószínűleg igen kevés aktív készítményt tartalmaz és bizonyára nagyszámú hatástalant is. A laboratóriumi állatokkal végzett standard próbákkal különbséget teszünk a pozitívok és negatívok között és (éppúgy, mint a varietás szelekciónál) célszerű szukcesszív lépcsőket alkalmazhatunk. Ilyenkor célunk az, hogy az utolsó lépésben is megtartott készítmények (melyeket azután természetesen sokkal komolyabb próbáknak

kell alávetni) a lehető legnagyobb hányadát tartalmazzák az aktív készítményeknek. Azonban ezen cél elérésekor tekintetbe kell venni a szelekciós program árát, amit elsősorban a felhasználásra szánt állatok száma határoz meg.

О РЕНТАБЕЛЬНОСТИ ЭКСПЕРИМЕНТИРОВАНИЯ

D. J. FINNEY

Резюме

Задача статистика в наше время состоит не только в анализе результатов экспериментов, но и в оказании помощи при планировании экспериментов. Необходимо сознавать, что планирование экспериментов в узком смысле (design) есть лишь одна из точек зрения планирования экспериментов в более широком смысле и нельзя допустить, чтобы участие статистика ограничивалось бы суженными комбинаторными проблемами планирования. Для того, чтобы усилия экспериментатора были наиболее плодотворными, необходимо тесное сотрудничество между статистиками и экспериментирующим ученым.

В докладе речь шла как о внутренней, так и о внешней экономии экспериментирования, стало быть о том, как следует планировать эксперимент так, чтобы использовать имеющиеся время, материал, работу с наибольшей выгодой, далее о том, какими возможностями надо располагать, чтобы эксперимент наилучшим образом служил интересам общества. В основных исследованиях часто имеет смысл исследовать лишь внутреннюю экономию, но при многих технологических экспериментах необходимо принимать во внимание и внешние факты. Во второй главе было приведено несколько особых вопросов, которые необходимо обсудить.

В главах 4—7 говорится о теории и практике некоторых специальных проблем планирования. Сначала говорилось об определении оптимального числа исследуемых факторов. По мнению автора и качественно и количественно наиболее выгодно одновременно вовлекать в эксперимент так много факторов, сколько позволяет размер эксперимента. Вторая группа вопросов — проблема возможности конкомитантной информации, относящейся к делянкам и другим экспериментальным единицам. Относительная выгода этого выявляется тогда, когда это применяется к образовыванию однородных блоков или когда результаты исправляются с помощью ковариантного анализа с конкомитантами.

Третья проблема говорит о том, как можно определить необходимое количество экспериментов для определения данного фактора. Один из видов этого имеет место, когда надо определить оптимальное число этих аналогичных экспериментов в таком случае, когда наша цель заключается в определении наилучшего уровня некоторого фактора для практического применения (например, количества минеральных удобрений на единицу площади посева). Это математически значительно проще, чем определение оптимального количества таких экспериментов, на основании которых нужно практически решить, какой из альтернативных и резко различных методов выбрать (например, две различные диеты). Доклад касался и этой проблемы

второго рода для случаев, когда имеется некоторая предварительная информация и экспериментатору на основании имеющихся данных нужно либо немедленно решать либо определить количество экспериментов, необходимых для решения.

Наконец, автор на двух примерах иллюстрирует вопрос об оптимальном планировании экспериментов, состоящих из последовательных ступеней. Один из них занимается селекцией вариантов выращиваемых растений, когда в последовательных сезонах производятся эксперименты на обработанной земле и в каждом сезоне выбирается некоторая часть наилучших для продолжения эксперимента в следующем сезоне. Если начальное число вариантов, число удержанных и в конце последнего сезона и полная площадь экспериментирования даны, то необходимо определить соотношение селекционных фракций и соотношение площадей земли, необходимых для отдельных стадий. Очевидно «одинаковые фракции в каждом году и равномерное распределение всей площади» дают схему, которая обычно близка к оптимальной. Второй пример говорит о том, что фильтруется очень большое число химических препаратов для определения тех, которые имеют терапевтическую ценность. Начальная популяция, вероятно, содержит очень мало активных изделий и очень много недействительных. С помощью лабораторных стандартных проб, произведенных на животных, различаем положительные и отрицательные и (также, как и в случае вариегатной селекции) можем применять целесообразные последовательные ступени. В таком случае цель заключается в том, чтобы изделия, удержанные при последнем шаге (которые затем, естественно, нужно подвергнуть значительно более серьезным экспериментам) содержали бы как можно большую часть активных изделий. Однако, при достижении этой цели необходимо принять во внимание цену селекционной программы, которая определяется в первую очередь числом используемых животных.

ПЯТЬ ЛЕКЦИЙ О НЕКОТОРЫХ ВОПРОСАХ ТЕОРИИ ЧИСЕЛ И ТЕОРИИ ВЕРОЯТНОСТЕЙ¹

Ю. В. ЛИННИК²

ЛЕКЦИЯ I

ДИСПЕРСИОННЫЙ МЕТОД ДЛЯ РЕШЕНИЯ НЕКОТОРЫХ БИНАРНЫХ АДДИТИВНЫХ ЗАДАЧ

§ 1. Задачи бинарные и тернарные

Работа И. М. Виноградова, G. H. HARDY, J. E. LITTLEWOOD, S. РАМАНУЖАН и многих других авторов дали теории чисел общий метод для решения многих аддитивных задач. Для выявления его характерных черт рассмотрим в самом кратком виде схему доказательства теоремы Гольдбаха—Виноградова. Для асимптотического решения уравнения: $p_1 + p_2 + p_3 = N$ (N нечётно), вводится сумма $S(\vartheta) = \sum_{p \leq N} \exp 2\pi i \vartheta p$. Если $Q(N) = \mathcal{C}(p_1 + p_2 + p_3 = N)$ (в дальнейшем: $\mathcal{C}(\cdot)$ — число решений (\cdot)), то:

$$(1.1) \quad Q(N) = \int_0^1 (S(\vartheta))^3 \exp(-2\pi i \vartheta N) d\vartheta.$$

Из сегмента $[0, 1]$ выделяется множество \mathfrak{M} точек ϑ , “аномально хорошо” аппроксимируемых рациональными дробями (они образуют легко описываемую систему сегментов небольшой меры), и дополнительное множество \mathfrak{m} . Тогда:

$$(1.2) \quad Q(N) = \int_{\mathfrak{M}} + \int_{\mathfrak{m}}.$$

Далее производится асимптотический расчёт $\int_{\mathfrak{M}}$ с помощью теорем о распределении простых чисел в арифметической прогрессии, а $\int_{\mathfrak{m}}$ оценивается следующим весьма характерным образом:

$$(1.3) \quad \int_{\mathfrak{m}} |S(\vartheta)|^3 d\vartheta \leq \sup_{\vartheta \in \mathfrak{m}} |S(\vartheta)| \int_0^1 |S(\vartheta)|^2 d\vartheta = \sup_{\vartheta \in \mathfrak{m}} |S(\vartheta)| \pi(N),$$

где $\pi(N) \sim \frac{N}{\ln N}$ — число простых чисел $\leq N$. Достаточно теперь получить

¹ Эти лекции были читанные в Математическом Институте Венгерской Академии Наук в октябре 1959 г.

² Ленинградское Отделение Математического Института им. В. А. Стеклова Академии Наук СССР.

оценку: при $\vartheta \in \mathfrak{m}$, $|S(\vartheta)| \leq \frac{S(0)}{g(N)}$, где $g(N) \geq \ln^3 N$, чтобы вывести асимптотическое решение проблемы.

Мы видим, что в получении нужной оценки $\int_{\mathfrak{m}}$, основную долю «понижающей работы» совершает равенство Parseval-я, простое аналитическое свойство ряда Фурье для $S(\vartheta)$, и после этого остаётся (весьма трудная) задача получения уже небольшого понижения в оценке остающейся $S(\vartheta)$. Эта же схема годна для весьма общего уравнения вида:

$$(1.4) \quad \alpha + \beta + \gamma = N,$$

где α, β, γ пробегают систему чисел достаточно хорошо распределённых в арифметических прогрессиях; α, β — достаточно «густы» в сегменте $[1, N]$ (скажем, количество их там не менее $\frac{N}{(\ln N)^c}$; $c > 0$ — константа), а дробные

доли $\{\vartheta\gamma\}$ при $\vartheta \in \mathfrak{m}$ ведут себя достаточно хорошо (сумма $\sum_{\substack{\gamma \leq N \\ \gamma \in \mathfrak{m}}} \exp 2\pi i \vartheta \gamma$ имеет достаточно хорошую оценку). При этом последовательность чисел γ может быть и весьма редкой, например, $\sum_{\substack{\gamma \leq N \\ \gamma \in \mathfrak{m}}} 1$ может быть порядка какой-либо степени $\ln N$ (см. соображения работ [1] и [2]).

Уравнение (1.4) является типично *тернарным*, для его решения было важно, грубо говоря, что α и β образуют две «густые» последовательности, а γ может быть весьма редкой, хотя и обладающей некоторой «диофантовой гладкостью» (в понятном из вышеизложенного смысле). Наличие не менее трёх слагаемых, два из которых образуют «густые» последовательности, а остальные действуют в качестве необходимого дополнения — весьма существенно для проведения описанной схемы расчёта.

Рассмотрим теперь классическое уравнение:

$$(1.5) \quad x^2 + y^2 + z^2 + t^2 = N.$$

«Точное» решение этого уравнения, сразу дающее асимптотику, было указано ещё Ясви [3] свыше ста лет назад. Хотя в (1.5) четыре слагаемых, они не образуют «густой» последовательности; достаточная густота получается только при их объединении в пары, а тогда задача (1.5) превращается в бинарную.

Решение этой задачи нельзя провести указанным выше методом. Более общее бинарное уравнение

$$(1.6) \quad ax^2 + by^2 + cz^2 + dt^2 = N$$

было решено Н. Д. Клоостерман-ом [4] в 1926 году методом контурного интегрирования с добавлением ряда особых соображений, весьма существенно использующих свойства квадратов чисел, и в 1954 г. М. Эйслер-ом [5] совершенно иным методом. Классическое уравнение $x^2 + y^2 + z^2 = N$, решенное К. Ф. Гаусс-ом также надо считать бинарным (по поводу общего уравнения $f(x, y, z) = N$; f — тернарная квадратичная форма см. [6], [7], [8]).

Уравнение HARDY—LITTLEWOOD-а (1923 г., см. [9])

$$(1.7) \quad p + \xi^2 + \eta^2 = N$$

также является с этой точки зрения бинарной задачей и решение эго методом контурного интегрирования не удаётся. В лекции 2 будет кратко изложено его решение «дисперсионным методом».

§ 2. Дисперсионный метод

Пусть $\{\varphi\}$ какая-либо последовательность натуральных чисел (допускается повторение); $\{D'\}$ пробегает без повторения какие-либо натуральные числа сегмента $[D_1, D_1 + D_2] = (D)$, $\{v\}$ независимо пробегает какие-либо натуральные числа сегмента $[v_0, v_0 + v'_0] = (v)$, а $n > (D_1 + D_2)(v_1 + v_2)$ — целое число.

Рассмотрим уравнение

$$(2.1) \quad n = \varphi + D'v,$$

представляющее бинарную задачу. Метод контурного интегрирования здесь не приводит к цели.

Пусть $U(m) = \sum_{\varphi=m} 1$. Число решений (2.1) задаётся суммой:

$$(2.2) \quad \sum_{D' \in (D)} \sum_{v \in (v)} U(n - D'v).$$

Пусть $D \in (D)$ какое-либо целое число. Рассмотрим при заданном D уравнение:

$$(2.3) \quad n = \varphi + Dv,$$

где $v \in (v)$. Пусть имеются какие-либо эвристические соображения (например, эвристическое применение метода контурного интегрирования) позволяющие предполагать, что ожидаемое число решений (2.3) имеет асимптотическое выражение

$$(2.4) \quad A(n, D).$$

Выражение:

$$(2.5) \quad V' = \sum_{D' \in (D)} \left(\sum_{v \in (v)} U(n - D'v) - A(n, D') \right)^2$$

будем называть дисперсией для данной проблемы при предполагаемой асимптотике $A(n, D)$. Это понятие является основным для дисперсионного метода.

Обратим внимание на то, что (2.5) представляет собой *двойную сумму*. Следуя основной идее метода И. М. Виноградова по оценке двойных сумм, мы можем только увеличить выражение V' , если в сумме (2.5) начнём суммировать подряд по всем целым значениям $D \in (D)$:

$$(2.6) \quad V' \leq V = \sum_{D_1 \leq D \leq D_1 + D_2} \left(\sum_{v \in (v)} U(n - Dv) - A(n, D) \right)^2.$$

Мы могли бы теперь заменить эвристическое нахождение нахождением его по методу наименьших квадратов, т. е. так, чтобы минимизировать V .

Разворачивая V , имеем: $V = V_1 - 2V_2 + V_3$, где

$$V_1 = \sum_{D_1 \leq D \leq D_1 + D_2} \left(\sum_{v \in (v)} U(n - Dv) \right)^2; \quad V_2 = \sum_{D_1 \leq D \leq D_1 + D_2} A(n, D) \sum_{v \in (v)} U(n - Dv);$$

$$V_3 = \sum_{D_1 \leq D \leq D_1 + D_2} (A(n, D))^2.$$

Сумма V_3 должна находиться путём несложного асимптотического подсчёта. Будем считать, что $v_0 \leq n^{1/2 - \varepsilon_0}$; $D_2 > n^{1/2 + \varepsilon_1}$ ($\varepsilon_0, \varepsilon_1$ — малые положительные константы), и что в таких условиях $\sum_{D_1 \leq D \leq D_1 + D_2} U(n - Dv)$, приводящаяся к числу решений сравнения $\varphi \equiv n \pmod{v}$ при некоторых ограничениях на величину φ допускает асимптотический расчёт.

Тогда можно сделать асимптотический расчёт V_2 . Расчёт V_1 является наиболее трудным. Выделяя из V_1 сумму $\sum_{D_1 \leq D \leq D_1 + D_2} \sum_{v \in (v)} (U(n - Dv))^2$, которую оцениваем грубо сверху, приходим к выводу, что оставшаяся часть V_1 равна совокупному числу решений системы уравнений:

$$(2.7) \quad \left. \begin{aligned} n - Dv_1 &= \varphi_2 \\ n - Dv_2 &= \varphi_1 \end{aligned} \right\},$$

где φ_1, φ_2 независимо пробегают числа последовательности $\{\varphi\}$, и числа $v_1 \neq v_2$ образуют пары (v_1, v_2) , где $v_i \in (v)$. Эти пары разбиваем на классы соответственно О. Н. Д. $(v_1, v_2) = \delta$. Элементарные выкладки показывают теперь, что при заданных (v_1, v_2) уравнения (2.7) равносильны одному уравнению:

$$(2.8) \quad v_1 \varphi_1 - v_2 \varphi_2 = n(v_1 - v_2)$$

при дополнительных условиях:

$$(2.9) \quad \varphi_1 \equiv \varphi_2 \equiv n \pmod{\delta}; \quad \frac{n - \varphi_1}{v_2} \in (D); \quad \frac{n - \varphi_2}{v_1} \in (D).$$

Если уравнение (2.8) удастся решить асимптотически с удовлетворительной погрешностью, то мы получим асимптотическое выражение для V_1 и вместе с тем, для V . При этом основное значение имеет отношение $\frac{V}{V_3}$.

Если оно достаточно мало, то при довольно общих условиях мы получаем асимптотическое решение бинарной задачи (2.3).

Будем считать далее, что $\{D'\}$ достаточно густо, именно, что количество чисел $\{D'\}$ не меньше $\frac{D_2}{(\ln n)^{K_1}}$ (далее $K_i > 0$ — большие константы).

Числа $v \in (v)$, напротив, могут образовывать весьма редкую последовательность.

Пусть теперь нам удалось удовлетворительно решить (2.8) и таким образом обнаружить, что:

$$(2.10) \quad V' \leq V \leq V_3 (\ln n)^{-K_2}.$$

Если K_2 достаточно велико, то отсюда немедленно следует разрешимость уравнения (2.3). В самом деле, если в (2.5) $U(n - Dv) = 0$ при всех

допустимых значениях D' и ν , то $V' = \sum_{D' \in (D)} (A(n, D'))^2$, а такое выражение не может быть настолько меньше V_3 , как этого требует неравенство (2.10). Более подробное рассмотрение с применением очевидного аналога неравенства Чебышева, даёт и асимптотику для (2.3). Сделаем некоторые дополнительные замечания. Областью изменения D' и ν в (2.3) у нас был прямоугольник. Можно заменить его гиперболической областью $D'\nu \leq n$, которую можно потом заменить суммой прямоугольников с допустимой погрешностью в числе решений. При этом, однако, существенно соблюдение условия: $\nu_0 \leq n^{1/2-\epsilon_0}$, т. е. числа ν не должны быть слишком большими. Далее, если ν пробегает какую-либо подпоследовательность простых чисел, или их степеней, то $\delta = 1$, и первое из условий (2.9) отпадает, что весьма удобно при решении (2.8). Ввиду этого, выгодно при решении уравнения (2.1) пытаться отщепить от ν простой множитель ν' и заменить $D'\nu$ на $D' \frac{\nu}{\nu'} \nu' = D''\nu'$,

при простом ν' , если только этот простой множитель не получается слишком малым. Далее, очевидное видоизменение метода позволит решить и уравнение вида $\varphi - D'\nu = m$, например: $\varphi - D'\nu = 2$ при $|\varphi| \leq n$; $n \rightarrow \infty$.

Далее, из наших рассуждений видно, что D' можно считать и имеющим повторение, но при условии, что среднее квадратичное число этих повторений не слишком велико.

§ 3. Случай последовательности квадратичных форм

Особенно хорошо поддаётся исследованию случай, когда числа φ пробегают (с соответствующим числом повторений) последовательность значений бинарных квадратичных форм: $\varphi = ax^2 + bxy + cy^2 = \varphi(x, y)$. При этом допускается любые целочисленные формы — $\varphi(x, y) = x^2 + y^2$; $\varphi(x, y) = x^2 - 2y^2$; $\varphi(x, y) = xy$ и т. д. Вычисление $A(n, D)$ и расчёт V_3 и V_2 не представляет тогда принципиальных трудностей. Уравнение (2.8) для расчёта V_1 принимает вид:

$$(3.1) \quad \nu_1 \varphi_1(x, y) - \nu_2 \varphi_2(x', y') = n(\nu_1 - \nu_2)$$

в условиях (2.9). Слева здесь стоит при данных ν_1, ν_2 неопределённая кватернарная квадратичная форма, т. е. получается обобщение задачи Клоостерман-а (1.6). Характерна большая величина коэффициентов ν_i . Подробное рассмотрение уравнения (3.1) на основе метода Клоостерман-а [4] и новейших оценок сумм Клоостерман-а (см. A. Weil, 1948 [10]; L. Carlitz, S. Uchiyama, 1957 [11]) показывает, что в общем случае уравнение (3.1) разрешимо с удовлетворительной асимптотикой при $\nu_i \leq \nu_0 = n^{1/10}$.

Если же $\varphi(x, y)$ — одноклассные квадратичные формы, т. е. отвечают квадратичному полю с одним классом идеалов, то решение (3.1) сравнительно просто, и удаётся уже при $\nu_i \leq \nu_0 = n^{1/6} \exp(-\sqrt{\ln n})$. Таковы случаи: $\varphi(x, y) = x^2 + y^2$; $\varphi(x, y) = x^2 \pm 2y^2$; $\varphi(x, y) = xy$ и некоторые другие. Это обстоятельство играет большую роль при асимптотическом решении уравнения Навру—Littlewood-а (1.7). Таким образом, мы имеем все средства к решению уравнений вида:

$$(3.2) \quad \varphi(x, y) + uv = n,$$

где u пробегает любую достаточно густую систему чисел (с не слишком многими повторениями), а v независимо пробегает систему чисел $v \leq \frac{n}{u}$ (лишь границы изменения v могут зависеть от u). Эта система может быть весьма редкой; нужно однако, чтобы было: $\exp(\ln n)^{e_2} \leq v \leq n^{1/6} \exp(-\sqrt{\ln n})$ (для $\varphi(x, y)$ одноклассной); $\exp(\ln n)^{e_2} \leq v \leq n^{1/10}$ (для общего случая $\varphi(x, y)$). Кроме того, числа v должны быть сравнительно хорошо распределены в арифметических прогрессиях с малой разностью. Это нужно для асимптотического расчёта V_2 .

Применение некоторого видоизменения решета Эратосфена и уравнению (3.2) позволяет решать и уравнение:

$$(3.3) \quad \varphi(x, y) + p = n.$$

ЛИТЕРАТУРА

- [1] Линник, Ю. В.: „Простые числа и степени двойки.” *Труды Математического Института им. Стеклова. АН СССР* **38** (1951) 152—169.
- [2] Линник, Ю. В.: „Складывание простых чисел со степенями одного и того же числа.” *Мат. Сборник* **32** (1953) 3—60.
- [3] Гурвиц, А.: *Теория аналитических и эллиптических функций*. Г. Т. Т. И., 1933, стр. 273.
- [4] Kloosterman, H. D.: „On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$.” *Acta Math.* **49** (1926) 407—464.
- [5] Eichler, M.: „Quaternäre quadratische Formen und die Riemannsche Vermutung für die Kongruenzzetafunktion.” *Archiv f. Math.* **5** (1954) 355—366.
- [6] Линник, Ю. В.: „Кватернионы и числа Кэли; некоторые приложения арифметических кватернионов.” *Усп. Мат. Наук* **4** (1949) **5** (33) 49—98.
- [7] Линник, Ю. В.—Малышев, А. В.: „Приложения арифметики кватернионов к теории тернарных квадратичных форм и к разложению чисел на кубы.” *Усп. Мат. Наук.* **8** (1953) **5** (57) 3—71.
- [8] Малышев, А. В.: „Асимптотическое распределение целых точек на некоторых эллипсоидах.” *Известия АН СССР, сер. матем.* **21** (1957) 457—500.
- [9] Hardy, G. H.—Littlewood, J. E.: „Some problems of partition numerorum, III.” *Acta Math.* **44** (1923) 1—70.
- [10] Weil, A.: „On some exponential sums.” *Proc. Nat. Acad. U.S.A.* **34** (1948) 204—207.
- [11] Carlitz, L.—Uchiyama, S.: „Bounds for exponential sums.” *Duke Math. J.* **24** (1957) N 1, 37—41.

ЛЕКЦИЯ 2

ПРИМЕНЕНИЕ ДИСПЕРСИОННОГО МЕТОДА. ПРОБЛЕМА ДЕЛИТЕЛЕЙ. ПРОБЛЕМА HARDY—LITTLEWOOD-A. ДРУГИЕ ПРОБЛЕМЫ

§ 1. Проблема делителей

Пусть $\tau_k(m)$ есть $\chi(x_1 x_2 \dots x_k = m)$ (как и в лекции 1, $\chi(\cdot)$ — число решений (\cdot)). Многие авторы, главным образом, английские, изучали асимптотическое поведение сумм вида:

$$(1.1) \quad \sum_{n \leq x} \tau_{k_1}(n) \tau_k(n + l); \quad \sum_{v \leq n} \tau_{k_1}(v) \tau_k(n - v).$$

См. напр. А. INGHAM [1]; Т. ESTERMANN [2]; Е. TITCHMARSH [3]; С. HOOLEY [4]. Были выведены асимптотические формулы для случаев $k_1 = 2$; $k = 2, 3$. Случай $k_1 = 2$, $k = 3$ был изучен С. HOOLEY (1957) в работе [4]. Дальнейшее продвижение на основе применяемых в указанных работах методах не удаётся.

Дисперсионный метод позволяет вывести асимптотику (1.1) для $k_1 = 2$ и любом целом $k = 2, 3, 4, \dots$. Мы сформулируем теоремы получающиеся при $l = 1$ (другие случаи отличаются от данного лишь более сложной записью).

Теорема 1. При $n \rightarrow \infty$, $k \geq 2$ имеем:

$$(1.2) \quad \sum_{m \leq n} \tau_2(m) \tau_k(m+1) \sim k! C_{k-1} S_k n (\ln n)^{k+1},$$

где

$$(1.3) \quad S_k = \sum_{m=1}^{\infty} \frac{\mu(m)}{m^2} \prod_{p|n} \left(p \left(1 - \left(1 - \frac{1}{p} \right)^{k-1} \right) \right);$$

$$C_{k-1} = \lim_{Y \rightarrow \infty} \frac{1}{(\ln Y)^{k-1}} \int \dots \int \frac{dy_1 \dots dy_{k-1}}{y_1 y_2 \dots y_{k-1}} \cdot \frac{Y}{y_1 y_2 \dots y_{k-1}},$$

$$1 \leq y_1 \leq y_2 \leq \dots \leq y_{k-1} \leq \frac{Y}{y_1 y_2 \dots y_{k-1}}.$$

Разумеется, константу C_{k-1} можно подсчитать явно и выразить через элементарные функции от k . При $l \neq 1$, выражение (1.3) значительно усложняется.

Имеет место и асимптотическое разложение вида:

$$(1.4) \quad \sum_{m \leq n} \tau_2(m) \tau_k(m+1) = n P_{k+1}(\ln n) + O(n(\ln n)^{\alpha_0}),$$

где $\alpha_0 > 0$ сколь угодно малая константа; P_{k+1} — полином степени $k+1$. Дадим схему применения дисперсионного метода к выводу асимптотики (1.1). Ограничимся случаем $l = 1$. Тогда дело сводится к уравнению:

$$(1.5) \quad x_1 x_2 \dots x_k - xy = 1; \quad xy = n.$$

Не нарушая общности, можем считать: $x_1 \geq x_2 \geq \dots \geq \dots x_k$. Полагая: $x_1 x_2 \dots x_{k-1} = D'$; $x_k = v$; $xy = \varphi$ приходим к уравнению вида:

$$(1.6) \quad D' v - \varphi = 1; \quad \varphi \leq n$$

(см. лекцию 1). Надлежит ещё обеспечить условие:

$$(1.7) \quad \exp(\ln n)^{\epsilon_1} \leq v \leq n^{1/6} \exp(-\sqrt{\ln n})$$

(см. лекцию 1). Для этого пытаемся представить $D'v$ в виде $D''v'$, где v' подчинено неравенствам (1.7).

Если этого нельзя сделать, то в (1.5) число $x_1 \dots x_k$, либо имеет все простые множители больше $n^{1/6} \exp(-\sqrt{\ln n})$, либо все меньше $\exp(\ln n)^{\epsilon_2}$. Для таких чисел $x_1 \dots x_k$ можно дать грубую оценку сверху числа решений (1.5). Для остальных можно применять дисперсионный метод, разбивая область изменения D'' и v' (которые переобозначим в D' и v) на удобные сегменты (мы не будем останавливаться на несущественных видоизменениях, которые потребуются внести).

Число $A(n, D)$ здесь надлежит выбрать в виде :

$$(1.8) \quad A(n, D) = \frac{n \ln n}{D^2} \varphi(D) \left(1 - \frac{1}{\ln n} \left(1 + 2 \frac{\Gamma'(1)}{\Gamma(1)} + 2 \sum_{p|D} \frac{\ln p}{p-1} \right) \right).$$

После этого оценивается дисперсия V . Если, например, $D_1 \leq D \leq D_1 + D_2$, где $D_1 \geq n^{5/6+\varepsilon_2}$, $D_2 = D_1 n^{-\varepsilon_3}$, то после соответствующих подсчётов получается дисперсия для данной зоны изменения D (зональная дисперсия) :

$$(1.9) \quad V = \sum_{D_1 \leq D \leq D_1 + D_2} \left(\sum_{\substack{m \equiv -1 \pmod{D} \\ m \leq n}} \tau_2(m) - A(n, D) \right)^2 = O \left(\frac{n^{2-\varepsilon_4}}{D_1^2} D_2 \right).$$

Применения очевидного аналога неравенства Чебышева приводит после этого к выводу асимптотики числа решений (1.5) для данной зоны изменения D'' . Собирая по зонам, после довольно кропотливых подсчётов приходим к выводу (1.2) и (1.4). Подробнее см. [5]. Надо отметить, что в отношении упрощения вывода выгоднее выбирать v' в разложении $x_1 x_2 \dots x_k = D'' v'$ простым числом. При этом, однако, $A(n, D)$ получит вид отличный от (1.8).

При $k \geq 3$, $k_1 \geq 3$ трактовка (1.1) дисперсионным методом не удаётся ; получается, однако, ряд любопытных условных теорем.

§ 2. Обобщение теоремы Kloosterman-а

Уравнение KLOOSTERMAN-а (см. лекцию 1)

$$(2.1) \quad ax^2 + by^2 + cz^2 + dt^2 = n$$

можно записать в виде :

$$(2.2) \quad N(\alpha_1) + N(\alpha_2) = n,$$

где α_1 и α_2 — идеалы квадратичных полей — соответственно $k(\sqrt{-ab})$ и $k(\sqrt{-cd})$, отвечающие так называемым амбиговым классам. Методом KLOOSTERMAN-а можно асимптотически решить и уравнение вида (2.2), где идеалы α_1 и α_2 берутся из заданных классов идеалов заданных квадратичных полей $k(\sqrt{d_1})$ и $k(\sqrt{d_2})$. Если поставить ту же задачу для идеалов из заданных классов других полей (не квадратичных), то метод KLOOSTERMAN-а перестаёт действовать.

Применение дисперсионного метода позволит асимптотически решить уравнение

$$(2.3) \quad N(\alpha_1) + N(\alpha_2) = n,$$

когда α_1 — идеал из заданного класса идеалов данного квадратичного поля k_1 , а α_2 — идеал из заданного класса идеалов любого заданного поля k_2 . Мы не будем приводить здесь довольно громоздкой асимптотической формулы, и поясним лишь схему решения. Так как k_1 — квадратичное поле, то $N(\alpha_1) = \varphi(x, y)$ где φ — квадратичная форма (при некоторых ограничениях на x, y). Далее, полагаем $N(\alpha_2) = D^v$, где v — простое число под условием

(1.7). Если такого разложения нет, то даём грубую оценку числа подобных решений (2.3); их будет относительно мало. Так приходим к уравнению:

$$(2.4) \quad \varphi(x, y) + D'v = n.$$

Разбиваем решение (2.4) на классы, где v — простое число норма идеала заданного класса, а D' пробегает нормы идеалов соответствующего определённого класса, так что $D'v = N(a_2)$; a_2 из предписанного класса. Далее уже к (2.4) можно применять дисперсионный метод и вывести асимптотику (2.3).

§ 3. Аналоги уравнения Hardy—Littlewood-a

Сравнительно просто применение дисперсионного метода к уравнению

$$(3.1) \quad n = \xi^2 + \eta^2 + p_1 p_2,$$

где p_1 пробегает простые числа сегмента $[1, N_1]$, p_2 — простые числа сегмента $[1, N_2]$; здесь $N_1 = n^{1-a}$; $N_2 = n^a$, a — любая константа под условием:

$$(3.2) \quad 0 < a < \frac{1}{6}.$$

Число решений (3.1) обозначим через $Q_1(n)$. Имеет место теорема:

Теорема.

$$(3.3) \quad Q_1(n) = \pi A_0 Li(N_1) Li(N_2) \prod_{p|n} \frac{(p-1)(p-\chi_4(p))}{p^2 - p + \chi_4(p)} + R_1(n),$$

где:

$$(3.4) \quad A_0 = \prod_p \left(1 + \frac{\chi_4(p)}{p(p-1)} \right); \quad Li(x) = \int_2^x \frac{dx}{\ln x}; \quad R_1(n) = O\left(\frac{n}{(\ln n)^c}\right)$$

и с сколь угодно большая константа.

Укажем схему применения дисперсионного метода. Здесь берём: $\varphi = \xi^2 + \eta^2$; $p_1 = D'$; $p_2 = v$. $A(n, D)$ имеет довольно сложный вид (легко находимый эвристическим применением решета Эратосфена: при $p_i | n$ пусть $p_i^{A_i} || D$, $p_i^{Q_i} || n$. Тогда:

$$(3.5) \quad A(n, D) = \pi Li(N_2) A_0 \prod_{\substack{p|D \\ p \nmid n}} (1 + \xi_p) \prod_{p_i|n} (1 + \xi_{p_i}) (1 + \eta(p_i, A_i)),$$

$$\text{где: } 1 + \xi_p = \frac{(p-1)(p-\chi_4(p))}{p^2 - p - \chi_4(p)}, \text{ и:}$$

$$1 + \eta(p_i, A_i) = \begin{cases} 1 + \chi_4(p_i) + \dots + \chi_4^{A_i}(p_i) & \text{при } A_i < Q_i; \\ 1 + \chi_4(p_i) + \dots + \chi_4^{Q_i}(p_i) & \text{при } A_i > Q_i; \\ 1 + \chi_4(p_i) + \dots + \chi_4^{Q_i}(p_i) + \frac{\chi_4^{Q_i+1}(p_i) p_i}{(p_i-1)(p_i-\chi_4(p_i))} & \text{при } A_i = Q_i. \end{cases}$$

Оценка дисперсии V требует несложных, но довольно кропотливых подсчётов; получается:

$$(3.6) \quad V = O(N_1 N_2^2 (\ln n)^{-c})$$

при сколь угодно большой константе c . Далее очевидный аналог неравенства Чебышева приводит к (3.3).

Если вместо простых чисел p_2 брать их степени p_2^a в том же сегменте $[1, N_2]$ при любом заданном $a (a = 1, 2, 3, \dots)$, т. е. рассматривать уравнение:

$$(3.7) \quad n = \xi^2 + \eta^2 + p_1 p_2^a,$$

то соответствующие выводы почти не отличаются от случая $a = 1$; если $Q_a(n)$ — число решений (3.7), то:

Теорема 3.

$$(3.8) \quad Q_a(n) = \pi A_0 Li(N_1) Li(N_2^{1/a}) \prod_{p|n} \frac{(p-1)(p-\chi_4(p))}{p^2 - p + \chi_4(p)} (1 + o(n)),$$

где $o(n) = O((\ln n)^{-c})$; $c > 0$ — сколь угодно большая константа.

Вместо чисел p_2^a в (3.7) удаётся поставить чрезвычайно разряжённую последовательность. Если рассмотреть уравнение:

$$(3.9) \quad n = \xi^2 + \eta^2 + p d^m,$$

где d, a фиксированы, и $d^m \leq N_2$; m переменна, то более тонкое применение дисперсионного метода приводит к асимптотике числа решений $Q'(n)$ уравнения (3.9):

Теорема 4.

При $(d, 2n) = 1$ имеем:

$$(3.10) \quad Q'(n) \sim \frac{\pi A_0}{(\ln d)^{1/a}} \frac{N_1}{\ln N_1} (\ln N_2)^{1/a} \prod_{p|nd} \frac{(p-1)(p-\chi_4(p))}{p^2 - p + \chi_4(p)}.$$

Если рассматривать уравнение (3.1), не накладывая никаких ограничений на p_1, p_2 , то вывод асимптотики затрудняется, и остаточный член резко ухудшается (за счёт малых значений $v = p_2$). Всё же без особых усложнений можно найти асимптотику для числа решений $Q''(n)$ (3.1) в этих условиях:

Теорема 5.

$$(3.11) \quad Q''(n) \sim \pi A_0 n \frac{\ln \ln n}{\ln n} \prod_{p|n} \frac{(p-1)(p-\chi_4(p))}{p^2 - p + \chi_4(p)}.$$

Сделаем два дополнительных замечания. Если вместо условия (3.2) поставить более жёсткое условие: $0 < \alpha < \frac{1}{10}$, то можно вывести асимптотические формулы, аналогичные теоремам 2, 3, 4, 5 при замене в соответ-

ствующих уравнениях $\xi^2 + \eta^2$ на любую целочисленную форму $\varphi(x, y) = ax^2 + bxy + cy^2$. Аналог теоремы 5, т. е. асимптотика уравнения :

$$(3.12) \quad n = ax^2 + bxy + cy^2 + p_1 p_2$$

может быть выведена без всяких условий.

Далее, как следует из соображений С. НООЛЕУ [6], (1957), уравнение HARDY—LITTLEWOOD-а $n = p + \xi^2 + \eta^2$, и теорема 5, могут быть выведены условно, из расширенной гипотезы РИЕМАНН-а. Надо отметить, что теоремы 2, 3 и 4 подобным образом из гипотезы РИЕМАНН-а не следуют, и потому в некотором смысле могут считаться более глубокими, чем уравнение HARDY—LITTLEWOOD-а. Соображения С. НООЛЕУ в [5] кроме того, неприменимы для случая общих форм а лишь для случая одноклассных форм, для дисперсионного метода такое различие не очень существенно.

§ 4. Уравнение Hardy—Littlewood-а

В применении к уравнению HARDY—LITTLEWOOD-а

$$(4.1) \quad n = p + \xi^2 + \eta^2$$

дисперсионный метод требует подготовительных преобразований. Здесь имеет место теорема для числа решений $Q(n)$ уравнения (4.1)

Теорема 6.

$$(4.2) \quad Q(n) = \pi \frac{n}{\ln n} \prod_p \left(1 + \frac{\chi_4(p)}{p(p-1)} \right) \prod_{p|n} \frac{(p-1)(p-\chi_4(p))}{p^2 - p + \chi_4(p)} + O(n(\ln n)^{-1.028}).$$

Формула (4.2) указали HARDY и LITTLEWOOD эвристически и без указания остаточного члена в 1923 г. в мемуаре [7].³

Для сведения уравнения (4.1) к виду, удобному для применения дисперсионного метода, положим $P = \exp \left(\frac{\ln n \ln \ln \ln n}{K \ln \ln n} \right) K$ — большая константа); Ω_P — множество чисел имеющих все простые множители больше P .

Полагая $\zeta_P(s) = \prod_{p>P} (1 - p^{-s})^{-1} = 1 + T(s)$, так что $\ln \zeta_P(s) = T(s) - \frac{1}{2} (T(s))^2 + \frac{1}{3} (T(s))^3 - \dots$ сведём (4.1) на уравнение Y'_k вида :

$$(4.3) \quad n = x'_1 \dots x'_k + \xi^2 + \eta^2; \quad x'_i \in \Omega_P.$$

Именно, если $Q_k(n)$ — число решений (4.3), то из разложения $\ln \zeta_P(s)$ непосредственно выводим (см. также [8])

$$(4.4) \quad Q(n) = \sum_{k=1}^{r_1} \frac{(-1)^{k+1}}{k} Q_k(n) + O(n^{3/4}).$$

³ В моей заметке [8] имелся пробел в доказательстве, благодаря чему остаточный член получается много хуже, чем указанный в [8].

Здесь $r_1 \leq \frac{K \ln \ln n}{\ln \ln \ln n}$. Если теперь положим: $x'_1 x'_2 \dots x'_n = D'v$, где v — наименьший простой делитель $x'_1 \dots x'_n$, то имеем, очевидно:

$$(4.5) \quad n^{1/k} \geq v > P.$$

При $k \geq 7$; $k \leq r_1$, простые числа v удовлетворяют условиям (1.7). Ввиду этого по всем уравнениям Y'_k вида (4.3) при $k = 7, 8, \dots, r_1$, непосредственно применим дисперсионный метод, и они сравнительно несложно решаются с хорошей асимптотикой (малой погрешностью). Остаются уравнения Y'_k с $k = 6, 5, 4, 3, 2, 1$. При $k = 6$ имеем: $n^{1/6} \geq v > P$.

Если при этом не выполнено (1.7), то аккуратный подсчёт показывает, что Y'_k будет иметь при таких v сравнительно мало решений. Их можно отбросить, а случаи остальных v — трактовать дисперсионным методом. Тогда Y'_6 решается с удовлетворительной асимптотикой, и остаются $Y'_5, Y'_4, Y'_3, Y'_2, Y'_1$. Уравнения Y'_k при всяком фиксированном k и $n \rightarrow \infty$ могут быть грубо говоря, сведены на уравнение Y_k вида:

$$(4.6) \quad n = x_1 x_2 \dots x_k + \xi^2 + \eta^2,$$

где x_i пробегает числа подряд. Эти уравнения надо решать с погрешностью $O\left(\frac{n}{(\log n)^c}\right)$ при $c > 1$. Это удаётся тривиальным образом при $k = 1$ и $k = 2$, и на основе оценок от А. Вейля для сумм Клоостермана (см. лекцию I) при $k = 3$. Остаются Y_4 и Y_5 .

Уравнения Y_k ($k = 4, k = 5$) могут быть решены, если располагать достаточно точными законами распределения в прогрессиях для $x_1 x_2 x_3 \dots x_k \pmod{D}$, где $D \leq \sqrt{n} \exp(-(\ln n)^\epsilon)$ ($\epsilon > 0$ сколь угодно мало). Сочетание подобных законов с некоторыми соображениями С. Ноолеса из работы [5] приводит к удовлетворительному решению Y_k . Простое применение теории L -рядов Дирихле к выводу таких законов в прогрессиях приводит к необходимости оценки для $\sum_{\chi \pmod{D}} \left| L\left(\frac{1}{2} + ti, \chi\right) \right|^k$.

Для $k = 4$, с помощью особого вида укороченных функциональных уравнений L -рядов (уравнений типа Харди—Литтлвуда) удаётся показать, что при $k = 4$

$$(4.7) \quad \sum_{\chi \pmod{D}} \left| L\left(\frac{1}{2} + ti, \chi\right) \right|^k = O(D(|t| + 2) \ln^9 D(|t| + 2)),$$

что приводит к вполне удовлетворительному решению Y_4 и Y'_4 . Так как: $\left| L\left(\frac{1}{2} + ti, \chi\right) \right|^4 + 1 \geq \left| L\left(\frac{1}{2} + ti, \chi\right) \right|^k$ при $k < 4$, то (4.7) годно и для $k = 3, 2, 1$ вместо Y , что даёт, между прочим, новый способ решения для Y_3, Y_2, Y_1 и затем Y'_3, Y'_2, Y'_1 , не требующий оценок от А. Вейля.

Остаётся уравнение Y'_5 . Если выделить среди $x'_1 \dots x'_5 = D'v$ такие значения простого v , которые удовлетворяют (1.7), то при таком дополни-

тельном условии, Y'_5 легко решается дисперсионным методом. Теперь воспользуемся тем, что $Q_5(n)$ входит в формулу (4.4) в виде члена $+\frac{1}{5}Q_5(n)$.

Можно отбросить такие случаи разбиения $x'_1 \dots x'_5 = D'v$, когда $n^{1/5} \geq v > n^{1/6} \exp(-\sqrt{\ln n})$, причём без труда обнаруживается, что при этом теряется не более 2,1% асимптотики. Так приходим к неравенству: при $n > n_0$:

$$(4.8) \quad Q(n) > 0,979 \pi \frac{n}{\ln n} \prod_p \left(1 + \frac{\chi_4(p)}{p(p-1)} \right) \prod_{p|n} \frac{(p-1)(p-\chi_4(p))}{p^2 - p + \chi_4(p)}.$$

Это, однако, ещё не асимптотика. Чтобы получить последнюю, желательно доказать (4.7) при $k=5$. Это не удаётся. Но, к счастью, законы распределения $x_1 \dots x_5$ в прогрессия нужны не индивидуально, а лишь в среднем. Удаётся доказать «усреднение» (4.7) при $k=6$ (и, стало быть, $k=5$):

$$\sum_{\frac{D_1}{2} \leq D \leq D_1} \sum_{z \bmod D} \left| L \left(\frac{1}{2} + ti, \chi \right) \right|^6 = O(D_1^2(|t|+2) \exp(\ln D(|t|+2))^\varepsilon),$$

где $\varepsilon > 0$ сколь угодно мало. Этого достаточно для доказательства (4.2). Аналогично решается и уравнение $n = p + xy$, т. е. находится асимптотика для $\sum_{p \leq n} \tau(n-p)$.

§ 5. Обобщение проблемы Hardy—Littlewood-а. Дополнительные замечания

Если $\varphi(x, y) = ax^2 + bxy + cy^2$ — любая бинарная форма, то уравнение

$$(5.1) \quad n = p + \varphi(x, y)$$

также поддаётся решению. Однако асимптотическая формула пока не получается. Выходит только довольно хорошая оценка снизу для числа решений (5.1).

Встаёт вопрос, что получается если $\{\varphi\}$ образует последовательность отличную от значений квадратичной формы. Решение уравнения:

$$(5.2) \quad n = \varphi + D'v$$

будет сводиться к решению уравнения (2.8) лекции 1, т. е. уравнение

$$(5.3) \quad v_1(n - \varphi_1) = v_2(n - \varphi_2)$$

в условиях (2.9) лекции 1. Если, например, $\varphi = \xi^3 + \eta^3 + \zeta^3 + \theta^3$, то подобные расчёты удаются. Весьма интересный случай: $\varphi = \xi^3 + \eta^3$ даёт в (5.3) проблему типа Waring-а, решение которой не удаётся. Он связан с другой, нерешённой проблемой HARDY—LITTLEWOOD-а [6].

Приближение к бинарной проблеме Гольдбаха вида:

$$(5.4) \quad n = p_1 p_2 + p_3 p_4,$$

где $\{\varphi\} = p_1 p_2$ даёт в (5.3) уравнение: $v_1(n - p_1 p_2) = v_2(n - p'_1 p'_2)$; v_i —

простые. Это уравнение содержит больше простых параметров, чем (5.4). Однако, пути к его решению не видно.

Наконец, случаи: $\{\varphi\} = \{x_1 x_2 x_3\}$ и $\{\varphi\} = \{x_1 x_2 x_3 x_4\}$ приводят к нерешённым, но не безнадёжным задачам о представлении системы чисел системой квадратичных форм.

ЛИТЕРАТУРА

- [1] INGHAM, A. E.: „Some asymptotic formulae in the theory of numbers”. *J. London Math. Soc.* **2** (1927) 173—182.
- [2] ESTERMAN, T.: „On the representation of a number as a sum of two products.” *Proc. Lond. Math. Soc.* **2** (31), (1930) 123—133.
- [3] TITCHMARSH, E. S.: „Some problems in the analytic theory of numbers”. *Quart. J. Math.* **13** (1942) 129—152.
- [4] HOOLEY, C.: „An asymptotic formula in the theory of numbers”. *J. Lond. Math. Soc.* **3** (1957) 396—419.
- [5] Линник, Ю. В.: „Дисперсия делителей и сумм квадратов в прогрессиях и некоторые бинарные аддитивные задачи.” *ДАН СССР* **120** (1958) 960—962.
- [6] HOOLEY, C.: „On the representation of numbers as the sum of two squares and a prime”. *Acta Math.* **97** (1957) 189—210.
- [7] HARDY, H.—LITTLEWOOD, J. E.: „Some problems of partitio numerorum, III”. *Acta Math.* **44** (1923) 1—70.
- [8] Линник, Ю. В.: „Проблема Гарди—Литтлвуда о сложении простых чисел и двух квадратов.” *ДАН СССР* **124** (1959) 9—30.

ЛЕКЦИЯ 3

ЦЕЛЫЕ ТОЧКИ НА СФЕРЕ И ЦЕПИ МАРКОВА. АНАЛОГИ ЭРГОДИЧЕСКИХ ТЕОРЕМ ДЛЯ ЦЕЛОЧИСЛЕННЫХ МАТРИЦ

§ 1. Целые точки на многомерных и трёхмерных сферах

На предыдущих лекциях рассматривалось уравнение Н. Д. Клоостермана $ax^2 + by^2 + cz + dt^2 = m$, т. е. вопрос о числе целых точек на четырёхмерном эллипсоиде. При $a = b = c = d = 1$ получаем четырёхмерную сферу $x^2 + y^2 + z^2 + t^2 = m$. В связи с этим встаёт естественный вопрос о распределении целых точек на k -мерной сфере $S\varphi_k(m): x_1^2 + x_2^2 + \dots + x_k^2 = m$.

С точки зрения возможных приложений наиболее интересна трёхмерная сфера $S\varphi_3(m): x^2 + y^2 + z^2 = m$. Если в пространстве имеется кубическая решётка (скажем, какого-либо кристалла), то представляет интерес распределение в пространстве узлов решётки на заданном расстоянии от данного. Притом естественно ожидать, что при больших значениях m , если целых точек на $S\varphi_3(m)$ много, то они будут распределены там асимптотически равномерно.

При $k \geq 5$ вопрос об асимптотическом распределении целых точек на $S\varphi_k(m)$ может быть исследован методом И. М. Виноградова—HARDY—LITTLEWOODA. При $k = 4$ этот вопрос требует применения метода Клоостермана, причём возникают значительные асимптотические трудности. Вопрос решён А. В. Малышевым [1], [2]; им доказана асимптотическая равномерность распределения точек на $S\varphi_4(m)$.

Случай трёхмерной сферы потребовал применения особого матричного аппарата, разработанного в ряде работ иной и А. В. Малышевым [3], [4], [5]. На основании этого аппарата, и с помощью некоторых соображений А. В. Малышева [6] вопрос о $S\varphi_3(m)$ в основном, решён мной в 1954 г. (см. [7], [8]). При решении этого вопроса выяснились интересные связи его с предельными теоремами для цепей Маркова и появляются аналоги эргодических теорем для целочисленных матриц.

Вопрос о самом существовании целых точек на $S\varphi_3(m)$ является глубоким и трудным. Из классических исследований К. Ф. Gauss-а известно, что они будут там существовать тогда (и, очевидно, только тогда), когда $m \neq 4^a(8b+7)$. Мы будем интересоваться только примитивными точками на $S\varphi_3(m)$; от них легко перейти ко всем точкам, количество их обозначим $H_0(m)$. Далее будем считать $m \equiv 1, 2, 3, 5, 6 \pmod{8}$, (что не нарушит общности рассмотрения). Число $H_0(m)$ может быть выражено с помощью классических формул Gauss-а через число классов идеалов поля $k(\sqrt{-m})$. Отсюда следует, согласно известной теореме С. Л. Siegel-я [9] (1935 г.), что при $m \rightarrow \infty$

$$(1.1) \quad \ln H_0(m) \sim \frac{1}{2} \ln m$$

так что число примитивных целых точек возрастает и можно говорить об их асимптотическом распределении на $S\varphi_3(m)$.

Рассмотрим сферу $S\varphi_3(1)$ концентрично с $S\varphi_3(m)$. Пусть на $S\varphi_3(1)$ дана замкнутая выпуклая область Γ_0 , ограниченная кусочно-гладким контуром, и $\omega(\Gamma_0)$ — телесный угол, под которым она видна из центра сферы, а Γ её проекция на $S\varphi_3(m)$. Имеет место теорема:

Теорема 1. (асимптотическая равномерность распределения). Пусть $q \geq 3$ простое число такое, что $\left(\frac{-m}{q}\right) = +1$, а Γ область указанного вида на $S\varphi_3(m)$ и $\omega(\Gamma)$ — телесный угол, под которым она видна из центра сферы. Тогда при $m \rightarrow \infty$ имеем

$$(1.2) \quad H_0(\Gamma) = \frac{\omega(\Gamma)}{4\pi} H_0(m) (1 + x_0(\Gamma_0, q, m))$$

$$(1.3) \quad H(\Gamma) = \frac{\omega(\Gamma)}{4\pi} H(m) (1 + x(\Gamma_0, q, m)).$$

Здесь $H_0(\Gamma)$ число примитивных точек внутри Γ , $H(\Gamma)$ — число всех целых точек внутри Γ ; $H(m) = H(S\varphi_3(m))$; x_0 и x стремятся к 0 при фиксированных Γ_0 и q , и $m \rightarrow \infty$.

В данной теореме участвует вспомогательное число q ; его наличие, повидимому, объясняется недостатками метода, а не существом дела. (иное положение имеется в излагаемых далее эргодических теоремах).

Не вводя постороннего числа q , можно доказать (1.2) и (1.3) (где уже не будет участвовать q) лишь условно, опираясь на некоторые недоказанные, хотя и очень слабые гипотезы об L — рядах DIRICHLET (см. [8], стр. 258 и [10]).

§ 2. Поток на примитивных точках $S\varphi_3(m)$. Формулировка эргодических теорем

Перейдём теперь к эргодическим свойствам множества примитивных целых точек на $S\varphi_3(m)$. Если $L(x, y, z)$ одна из таких точек, то отражая её в координатных плоскостях и поворачивая на 120° вокруг координатных «биссектрис», получим новые примитивные точки. Рассмотрим сферический треугольник на $S\varphi_3(m)$, ограниченный сечениями $S\varphi_3(m)$ плоскостями $z = 0$, $y - z = 0$, $x - z = 0$. Его будем называть основной областью Ω .

Пусть $q \geq 3$ простое число под условием $\left(\frac{-m}{q}\right) = +1$, а k — какое-либо натуральное число. Рассмотрим рациональные вращения и ортажения пространства, отвечающие числу q^k . Под этим будем понимать ортогональные матрицы $T = \left\| \frac{a_{ij}}{q^k} \right\|$, где a_{ij} целые числа, не все дроби сократимы, и $\det(T) = +1$. В результате вращения целая примитивная точка $L \in \Omega$ перейдёт в другую, вообще говоря, нецелую точку, которую мы обозначим $T \cdot L$. Если мы захотим выделить такие вращения T , которые переводят $L \in \Omega$ в $L' = TL \in \Omega$ и $L' \neq L$, то окажется, что таких вращений будет два и только два, в зависимости от выбора одного из двух решений сравнения $\xi^2 + m \equiv 0 \pmod{q^k}$, отличающихся только знаком. При этом точка $L' = T \cdot L$ будет примитивной вместе с L .

Выбирая одно из решений $\pm \xi_0$ указанного сравнения и фиксируя его, мы однозначно определим T и образ $L' = T \cdot L$. Полученную однозначную операцию вращения обозначим через T . Имеем: $T \cdot L = L'$.

Таким образом, на множестве a всех примитивных точек Ω операция T определяет поток. Повторение T r раз обозначим T^r . Спроектируем сферический треугольник Ω на $S\varphi_3(1)$ и обозначим полученный там треугольник через Ω_0 . Множество a спроектируется в a_0 и поток на нём индуцирует поток на a_0 . Этот поток обладает любопытными эргодическими свойствами.

Пусть $A_0 \subset \Omega_0$ односвязная область Ω_0 , ограниченная замкнутым кусочно-гладким контуром, $f_A(X)$ — характеристическая функция множества A_0 (равна 0 или 1); L_0 проекция на $S\varphi_3(1)$ $L \in a$; $\omega(A_0)$ — телесный угол, под которым видно множество A_0 .

Рассмотрим, при данном L_0 , точки: $L_0, TL_0, T^2L_0, \dots, T^{s-1}L_0$; пусть при этом $s \geq c_0 \ln m$ (c_0, C_i далее положительные константы).

Теорема 2 (эргодическая).

$$(2.1) \quad \frac{f_{A_0}(L_0) + f_{A_0}(TL_0) + \dots + f_{A_0}(T^{s-1}L_0)}{s} = \frac{6\omega(A_0)}{\pi} (1 + x(q^k, A_0, m))$$

для всех образов примитивных точек L_0 за возможным исключением $o(H_0(m))$, причём $x(q^k, A_0, m) \rightarrow 0$ при заданных q^k, A_0 и $m \rightarrow \infty$.

Поток T обладает и более сильным свойством перемешивания. Пусть \mathfrak{M}_0 какое-либо множество проекций точек из a на единичную сферу $S\varphi_3(1)$. Будем обозначать через $T'\mathfrak{M}_0$ множество, куда «перетекают» эти точки после r -кратного преобразования T ; $M(\mathfrak{M}_0)$ число точек \mathfrak{M}_0 ; $M(T'\mathfrak{M}_0 \cap A_0)$ — число точек множества $T'\mathfrak{M}_0$, лежащих в множестве A_0 .

Теорема 3 (о перемешивании). Пусть $l = 0, 1, 2, \dots$; $s \geq \theta_0 \ln m$ и $M(\mathbb{M}_0) > \theta_1 H_0(m)$ (θ_0, θ_1 — любые малые константы). Тогда для всех индексов p за возможным исключением $s o(1)$ таких индексов, имеем:

$$(2.2) \quad M(\mathbf{T}^l \mathbb{M}_0 \cap A_0) = \frac{6}{\pi} \omega(A_0) M(\mathbb{M}_0) (1 + x(q^k, A_0, \varepsilon_0, \varepsilon_1, m)),$$

где $x(q^k, A_0, \varepsilon_0, \varepsilon_1, m) \rightarrow 0$ при заданных $q^k, A_0, \varepsilon_0, \varepsilon_1$ и $m \rightarrow \infty$.

Важно заметить, что из теоремы 3 о перемешивании следует, как частный случай, теорема 1 об асимптотически равномерном распределении. Именно, примем за множество \mathbb{M} все множество α примитивных точек Ω . Очевидно, тогда: $\mathbf{T}^l \mathbb{M} = \mathbb{M}$ при любом значении l . Ввиду этого, из (1.5) следует:

$$(2.3) \quad M(\alpha \cap A_0) = \frac{6}{\pi} \omega(A_0) M(\alpha) \left(1 + x \left(q^k, A_0, \varepsilon_0, \frac{1}{24}, m \right) \right),$$

ибо $M(\alpha) = \frac{1}{24} H_0(m)$. Отсюда непосредственно следует теорема 1.

§ 3. Алгебраическая трактовка. Связь с теорией цепей Маркова

Доказательства теорем 1, 2, 3 удаются на основе трактовки вопроса с помощью кватернионов; это даёт также возможность довольно широкого обобщения. Мы ограничимся краткой формулировкой соответствующих понятий.

Уравнение $Sq_3(m): x^2 + y^2 + z^2 = m$ можно записать в виде $L^2 = -m$, где $L = xi + yj + zk$ вектор. Если ξ_0 избранное решение сравнения $\xi^2 + m \equiv 0 \pmod{q^k}$, то имеем: $\xi_0 + L = PX$, где P целый кватернион, $\mathcal{N}(P) = q^k$; X — целый кватернион. Выбор $L \in \Omega$ и поворотов внутри Ω фиксирует P среди 24 ассоциированных, и тогда операция потока $L' = TL$ алгебраическая изображается, как: $L' = P^{-1}LP$, $\xi_0 + L' = XP$. Пусть s большое число $s > c_0 \ln m$, притом такое, что $q^{ks} \leq \sqrt{m}$. Решая сравнение: $\xi^2 + m \equiv 0 \pmod{q^{ks}}$; $\xi \equiv \xi_0 \pmod{q}$, для каждого $L_\alpha \in \alpha$, получим равенство:

$$(3.1) \quad \xi + L_\alpha = P_{\alpha 1} P_{\alpha 2} \dots P_{\alpha s} \cdot V_\alpha, \quad \alpha = 1, 2, \dots, M(\alpha).$$

Несложно обнаружить, что эргодическая теорема 2 сводится, грубо говоря, к тому, что для всех значений $\alpha = 1, 2, \dots, M(\alpha)$, за возможным исключением $o(M(\alpha))$, относительные частоты встречи каждого возможного примитивного кватерниона $P_{\alpha\beta}$ в строчке (3.1) асимптотически одинаковы. Пусть $\pi_1, \pi_2, \dots, \pi_Q$ все примитивные кватернионы нормы q^k . Назовём их состояниями. Тогда строчка $P_{\alpha 1} P_{\alpha 2} \dots P_{\alpha s}$ формально образует "путь по состояниям" π_i . Если рассматривать возможные пути такого рода вне зависимости от равенства (3.1) и считать каждый из них равновероятным элементарным событием, то можно заметить, что получается схема однородных цепей Маркова. Именно, на первом месте возможно любое состояние π_{i1} . После π_{i1} возможны те и только те состояния π_{i2} , где $\pi_{i1} \pi_{i2}$ примитивно, после этого только такие состояния π_{i3} , где $\pi_{i2} \pi_{i3}$ примитивно и т. д., т. е. получается схема простой однородной цепи Маркова, где условные вероятности будущих состояний одинаковы, либо равны 0.

Пусть $\varrho(\pi_i)$ число появлений состояния π_i в строчке $P_{a1} \dots P_{as}$, так что $E\varrho(\pi_i) = \frac{s}{Q}$. Нас интересует оценка вероятности совмещения событий:

$$(3.2) \quad P \left\{ \left| \varrho(\pi_i) - \frac{s}{Q} \right| > \varepsilon s \right\} \quad i = 1, 2, \dots, Q,$$

где $\varepsilon > 0$ задано, а $s > c_0 \ln m$ возрастает. Эта оценка относится к теории больших уклонений для величин связанных в цепи Маркова (теория больших уклонений будет рассматриваться в лекции 5). Вероятность события (3.2) оценивается, как $O(e^{-\varepsilon_1 s})$; $\varepsilon_1 = \varepsilon_1(\varepsilon_0, Q) > 0$ (см. подробное изложение в [11]). Таким образом, возможных путей $P_{a1} \dots P_{as}$ в [3.4], где грубо говоря, не соблюдается эргодичность, будет относительно весьма мало. Особый и сравнительно трудный подсчёт (см. [3], [4], [5]) показывает, что и векторов L_a в (3.1), отвечающих таким «неэргодическим» путям, будет мало, т. е. $o(H_0(m))$ (это, разумеется, совершенно не очевидный, и довольно глубокий факт). Это и приводит к доказательству эргодической теоремы 2. Из неё несложно следует теорема о перемешивании 3; как было показано, частным случаем теоремы о перемешивании является теорема 1 об асимптотической равномерности.

Излагаемый метод, применявшийся к кватернионам $L^2 = -m$ и квадратичным полям $k(\sqrt{-m})$, может быть перенесен в известном смысле на целые матрицы любого порядка и любые конечные алгебраические расширения рационального поля. Таким образом возникают любопытные аналоги эргодических теорем в теории алгебраических чисел и целочисленных матриц, и новые асимптотические теоремы.

§ 4. Обобщения

Если уравнение: $L^2 = -m$ перепишем в виде:

$$(4.1) \quad L^2 = -mE \quad (E \text{ единица}), m > 0$$

и будем рассматривать его решения в целых матрицах 2 порядка, то вместо сферы $S\varphi_3(m)$ получим двуполостный гиперболоид: $ac - b^2 = m$. Проводя предыдущие рассуждения с заменой кватернионов на матрицы (см. подробное изложение в [5], [12]), получим любопытное усиление теоремы С. L. SIEGEL-я [9]: роль $S\varphi_3(m)$ будет играть область приведения Лагранжа $c \geq a \geq 2|b|$, на ней примитивные точки (a, b, c) , отвечающие примитивным бинарным квадратичным формам детерминанта $-m$ распределены равномерно в смысле мероопределения Лобачевского. Операция потока TL отвечает умножению классов идеалов, соответствующих L и одному и тому же идеалу \mathfrak{p} (см. подробнее в [13] и [14]).

Общая теория для довольно широкого класса алгебраических полей и матриц любого порядка (содержащих, например, как частный случай все Куммеров поля: $L^K = mE$) покоится на теореме СНАТЕЛЬЕРА—ШУРА (см. [13], [14]). Отдельные число арифметические стороны возникающей здесь картины отмечала О. ТАУССКУ [15], (повидимому, мало знакомой с работами А. СНАТЕЛЬЕРА и I. SCHUR-a). Однако, асимптотико-геометрическая и эргодическая сторона дела в излагаемом аспекте пока ещё не рассматривалась.

Отметим выясняющиеся по пути любопытные факты из арифметики целых квадратных матриц любого порядка r (см. [13], [14]).

Пусть L целая матрица, а Q целая матрица такая, что $\det(Q)$ свободен от квадратов.

Лемма. Для того, чтобы матрица $L' = Q^{-1}LQ$ была целой, необходимо и достаточно наличие равенства: $lE + L = QU$, где l — целое число, U целая матрица, E единица.

Эта лемма даёт возможность обобщения предыдущих теорем о сфере. Здесь операции потока $L' = TL = Q^{-1}LQ$ отвечает умножение на идеал $\alpha = (\det(Q), l + \vartheta)$, где ϑ один из корней минимального характеристического уравнения матрицы L .

С точки зрения асимптотики, представляет интерес следующая теорема. Мы сформулируем её (только для простоты) в случае Куммеровского поля :

$$(4.2) \quad L^K = mE.$$

Рассмотрим совокупность целых матриц L , удовлетворяющих (4.2) таких, что $L = O(m^{1/k})$. Изобразим их в виде точек в r^2 — мерном пространстве и рассмотрим все матрицы L , лежащие в круговом конусе раствора $\leq \varepsilon$.

Теорема 4. При достаточно малом фиксированном $\varepsilon > 0$, и $m \rightarrow \infty$ для любых двух матриц L, L' из нашего конуса возможен целый «поворот»

$$(4.3) \quad QLQ^{-1} = L',$$

где $Q = O(m^{1/2k})$.

Можно надеяться получить на пути развития аналогов эргодических теорем для целых матриц, в связи с теорией алгебраических чисел довольно глубокие асимптотические теоремы.

ЛИТЕРАТУРА

- [1] Малышев, А. В. : „О распределении целых точек на четырёхмерной сфере.” ДАН СССР 114 (1957) 5—8.
- [2] Малышев, А. В. : „О представлении целых чисел положительными квадратичными формами с 4 и более переменными, I.” Изв. АН СССР, серия мат. 3 (1959) 337—364.
- [3] Линник, Ю. В. : „Кватернионы и числа Кэли ; некоторые применения арифметики кватернионов.” УМН 4 (1949) 5 (33) 49—98.
- [4] Линник, Ю. В. — Малышев, А. В. : „Приложение арифметики кватернионов к теории тернарных квадратичных форм и к разложению чисел на кубы.” УМН 8 (1953) 5 (57) 3—71. поправка : 10 1955.
- [5] Линник, Ю. В. : „Асимптотическое распределение приведенных бинарных квадратичных форм в связи с геометрией Лобачевского.” I.: Вестн. ЛГУ № 2 (1955) 3—23 ; II.: № 5 (1955) 3—32 ; III. : № 8 (1955) 15—27.
- [6] Малышев, А. В. : „Асимптотический закон для представления чисел некоторыми положительными тернарными квадратичными формами.” ДАН СССР 93 (1953) ; 771—774.
- [7] Линник, Ю. В. : „Асимптотическое распределение целых точек на сфере.” ДАН СССР 96 (1954) 909—912.
- [8] Линник, Ю. В. : „Асимптотико-геометрические и эргодические свойства множества целых точек на сфере.” Мат. Сборник 43 (85) : 2 (1957) 257—276.
- [9] SIEGEL, C. L. : „Über die Klassenzahl quadratischer Zahlkörper”. Acta Arithm. 1 (1935) 83—86.

- [10] МАЛЫШЕВ, А. В.: „О связи теории распределения нулей L -рядов с арифметикой тернарных квадратичных форм.” *ДАН СССР* **122** (1958) 343—345.
- [11] Линник, Ю. В.: „Цепи Маркова в аналитической арифметике кватернионов и матриц.” *Вестник ЛГУ* № **13** (1956) 63—68.
- [12] LINNIK, JU. V.: „An application of matrices and of Lobatschevskian geometry to the theory of Dirichlet's real characters”. *Journ. Ind. Math. Soc.* **20** (1956) 37—45.
- [13] Линник, Ю. В.: „Ещё об аналогах эргодических теорем для мнимого квадратичного поля.” *ДАН СССР* **109** (1956) 694—696.
- [14] Линник, Ю. В.: „Некоторые применения неевклидовых геометрий к теории характеров Дирихле; аналоги эргодических теорем.” *Тр. III. Всес. Съезда Математиков* **3** (1959) 21—29.
- [15] TAUSKY, O.: „On a theorem of Latimer and Mac Duffee.” *Canadian J. Math.* **1** (1949) 300—303.

ЛЕКЦИЯ 4

О НЕКОТОРЫХ СВОЙСТВАХ БЕЗГРАНИЧНО ДЕЛИМЫХ ЗАКОНОВ

§ 1. Некоторые аналитические задачи теории вероятностей

Наиболее значительная часть современной теории вероятностей — носит асимптотический характер. Другие важные теоремы её содержат асимптотические условия с самой формулировке (напр. аксиомы, характеризующие процессы Poisson-а; условия непрерывности или иных функциональных свойств в вероятностью 1 тех или иных стохастических процессов и др.).

«Точные» высказывания теории вероятностей носят не столь значительный и разрозненный характер, и могут быть отнесены скорее к «вероятностному анализу» (хотя подобные определения — дело вкуса). Сюда относятся некоторые факты теории малых выборок, задачи о независимых и одинаково распределённых статистиках, «алгебры распределений» и некоторые другие.

Тем не менее, и такие факты полезны для теории вероятностей и математической статистики (особенно, теории малых выборок). Кроме того точным и простым формулировкам иногда присуще своеобразное изящество.

Остановимся на нескольких примерах этого рода.

1. Теорема Скитовича—DARMOIS [1].

Если две линейные формы: $L_1 = a_1 X_1 + \dots + a_n X_n$ и $L_2 = b_1 X_1 + \dots + b_n X_n$ от независимых случайных величин X_1, \dots, X_n независимы, то для тех индексов j , для которых $a_j b_j \neq 0$, величины X_j нормальны.

Эта теорема подверглась значительному обобщению в работе А. А. Зингера [2], где рассмотрен довольно широкий класс нелинейных независимых статистик (см. также [3]).

2. Приведём теорему (см. Ю. В. Линник [4]), примыкающую к направлению работы А. РЕНУИ [5]. Пусть x_1, \dots, x_n независимые случайные наблюдения. Рассматриваем две линейные статистики:

$$L_1 = a_1 x_1 + \dots + a_r x_r$$

$$L_2 = b_1 x_1 + \dots + b_r x_r$$

удовлетворяющие условию: $\sup(|a_1|, \dots, |a_r|) \neq \sup(|b_1|, \dots, |b_r|)$.
Для эквивалентности двух утверждений:

(А) наблюдения x_i принадлежат к нормальному типу распределений

(В) статистики L_1 и L_2 одинаково распределены — необходимо и достаточно выполнение следующих 5 условий.

Пусть $G(z) = |a_1|^z + \dots + |a_r|^z - |b_1|^z - \dots - |b_r|^z$ (z — комплексный аргумент). Должно быть:

- 1) $a_1 + a_2 + \dots + a_r = b_1 + b_2 + \dots + b_r$
- 2) $G(2) = 0$
- 3) Все положительные корни $G(z)$, являющиеся целыми числами, делящимися на 4, должны быть простыми корнями.
- 4) Все положительные корни $G(z)$, являющиеся целыми чётными числами $\equiv 2 \pmod{4}$ должны иметь кратность не выше 2, причём если есть такой двукратный корень, он должен быть единственным и максимальным из всех положительных корней $G(z)$.
- 5) Если $G(z)$ имеет положительный корень γ , не являющийся целым чётным числом, он должен быть единственным, максимальным из положительных корней, простым, и $\left[\frac{\gamma}{2} \right] \left(\text{целая часть } \frac{\gamma}{2} \right)$ должна быть нечётной.⁴

Эта теорема может быть приложена к сведению критериев нормальности на критерий однородности. Обобщение её на случай нескольких линейных статистик и нескольких типов наблюдений должно быть весьма любопытным.

3. Проблема Н. Свамёр-а. Н. Свамёр в своём известном обзоре [6] среди ряда проблем, поставил общую проблему разложений безгранично делимых законов. Сюда принадлежат известные отдельные результаты самого Н. Свамёр [7] — нормальный закон может быть разложен только на нормальные независимые компоненты, и Д. А. Райкова [8]: закон Poisson-а может быть разложен только на Пуассоновы компоненты.

В дальнейшем мы будем заниматься проблемой Н. Свамёр-а и её обобщениями.

§ 2. Разложения безгранично делимых законов

Логарифм характеристической функции (в дальнейшем х. ф.) безгранично делимого (в дальнейшем б. д.) закона даётся известной формулой:

$$(2.1) \quad \ln \varphi(t) = \beta it - \gamma t^2 + \int_{-\infty}^0 \left(e^{itu} - 1 - \frac{itu}{1+u^2} \right) dG_-(u) + \\ + \int_0^{\infty} \left(e^{itu} - 1 - \frac{itu}{1+u^2} \right) dG_+(u),$$

где β и $\gamma \geq 0$ реальны; $G_-(u)$ и $G_+(u)$ — «спектральные функции» неубывающие и такие, что $G_-(-\infty) = G_+(\infty) = 0$ и $\int_a^0 u^2 dG_-(u) + \int_0^a u^2 dG_+(u) < \infty$ для любого конечного a .

⁴ В реферате Chung на работу [4] замечено, что по его мнению, в § 52 имеется пробел. Проверка § 52 опровергла это.

Класс всех б. д. законов обозначим через I . Если б. д. закон $F \in I$ имеет компоненту F_1 (т. е. $F = F_1 * F_2$), где F_1 и F_2 — вероятностные интегральные законы, и $*$ знак композиции, то компоненты б. д. закона вообще говоря не обязаны быть б. д., т. е. $F_i \notin I$ вообще говоря. Но некоторые виды б. д. законов $F \in I$ могут иметь только б. д. компоненты.

Таковы, например, нормальный закон и закон Poisson-a. Класс б. д. законов, имеющих только б. д. компоненты, обозначим через I_0 . В дальнейшем будем заниматься задачей описания I_0 .

Мы будем здесь заниматься лишь случаем, когда б. д. закон F содержит нормальную компоненту ($\gamma > 0$). Имеет место теорема (см. Ю. В. Линник [9]).

Теорема 1. Для того, чтобы безгранично делимый закон F , имеющий гауссову компоненту, разлагаясь только на б. д. компоненты, необходимо, чтобы его спектральные функции $G_-(u)$ и $G_+(u)$ имели лишь конечное или счётное множество точек роста. При том, эти точки роста должны иметь вид:

$$(2.2) \quad \text{для} \quad G_+(u) : \dots k_{-2} k_{-1} \mu, k_{-1} \mu, \mu, \frac{\mu}{k_1}, \frac{\mu}{k_1 k_2}, \dots$$

$$(2.3) \quad \text{для} \quad G_-(u) : \dots -l_{-2} l_{-1} v, -l_{-1} v, -v, -\frac{v}{l_1}, -\frac{v}{l_1 l_2}, \dots$$

$$\text{где} \quad \dots k_{-2}, k_{-1}, k_1, k_2, \dots; \quad \dots l_{-2}, l_{-1}, l_1, l_2, \dots$$

какие-либо наборы натуральных чисел (допускаются повторения); $\mu \geq 0$ и $v \geq 0$ какие-либо числа.

Если множество точек роста спектральных функций ограничено, т. е. $dG_+(u) = dG_-(u) = 0$ при $|u| > K$, то это необходимое условие является и достаточным (далее будет указано и более широкое достаточное условие).

Подробное доказательство этой теоремы, основанное на методе перенала, теореме RALEY—WIEBENЯ на целых функциях и некоторых свойствах интеграла ЛЕБЕСГЕ-а изложено в [9]. Данное там изложение может быть значительно упрощено.

Возникает естественное предположение, что высказанное здесь необходимое условие (где весьма существенно, что $\gamma > 0$), совпадает с достаточным (где, к стати, это не существенно). Однако, при доказательстве этого предположения возникают большие трудности. Будет уместно их несколько проанализировать. При исследовании достаточных условий принадлежности б. д. $F \in I_0$ в указанной работе мной применяется аппарат целых функций. Но логарифм х. ф. F не обязан быть не только целой функцией, но и функцией продолжимой с реальной оси вообще. Он будет целой функцией, если спектр F ограничен, или вообще если «энергии высоких частот» $dG_+(u)$ и $dG_-(u)$ при большом $|u|$ достаточно быстро убывают.

Естественной идеей является осуществление следующего предельного перехода: при заданном б. д. F со счётным спектром указанного выше вида, заменяем F на последовательность законов $F_N (N \rightarrow \infty)$ таких, что $G_{N-}(u)$ и $G_{N+}(u)$ совпадают с $G_-(u)$ и $G_+(u)$ при $|u| \leq N$, и далее не имеют точек роста. По указанной выше теореме, все законы $F_N \in I_0$. Можно ли заключить отсюда, что и предельный закон $F \in I_0$? Решение данной задачи не удалось, хоть надо думать, что она решается положительно.

Пока удалось доказать достаточность положенного выше условия на спектр для случая, когда «энергия высоких частот» весьма мала. Именно, условие достаточно, если частоте $\mu_m > K$ и указанного выше вида отвечает «энергия» $dG_+(\mu_m) = \lambda_m$ такая, что $\ln \ln \frac{1}{\lambda_m} > C \mu_m^{1+\alpha}$, а частоте $-\nu_n < -K$ — энергия λ_{-n} такая, что $\ln \ln \frac{1}{\lambda_{-n}} > c \nu_n^{1+\alpha}$ (здесь K — большая константа; $\alpha > 0, c > 0$ — любые константы). Это условие, повидимому, можно значительно ослабить и довести до: $\ln \frac{1}{\lambda_m} < c \mu_m^{2+\alpha}$; $\ln \frac{1}{\lambda_{-n}} < c \nu_n^{2+\alpha}$, но можно ли двигаться далее, неясно.

Отметим также, что указанные выше достаточные условия действуют и при отсутствии гауссовой компоненты ($\gamma = 0$), но тогда явно не являются необходимыми.

Обратимся к необходимым условиям. Желательно сформулировать их для случая отсутствия гауссовой компоненты ($\gamma = 0$). Повидимому, конечность или счётность спектра во всяком случае должна быть необходимым условием, т. е. непрерывный спектр не может давать законы $F \in I_0$. В работе Н. Срамёр-а [10] (1949) для одного частного случая непрерывного спектра доказывается, что $F \notin I_0$, но общий случай непрерывного спектра ещё не разобран. Если указанное предположение верно, то можно надеяться на полное описание законов $F \in I_0$, без гауссовой компоненты и с не более, чем счётным спектром (частично это сделано ещё Д. А. Райковым в [8]).

Из остающихся в проблеме описания I_0 задач надо отметить ещё распространение теории на случай многомерных б. д. законов (описание I_0 в многомерном случае). Здесь почти ничего не известно.

Идя далее, мы можем рассмотреть случайные величины со значениями в пространстве Ванахса-а. Прежде всего; может представить интерес задача о разложении характеристических функционалов на характеристические функциональные множители. Пусть характеристический функционал Φ является аналитическим функционалом (см. Е. Ниле [11], стр. 89) и $\Phi = \Phi_1 \cdot \Phi_2$, где Φ_1 и Φ_2 характеристические функционалы.

Если наше пространство Ванахса-а — конечномерное векторное пространство, то Φ_1 и Φ_2 будут тоже аналитическими в той же области, что и Φ . Но будет ли так в общем случае? И как будут себя вести в отношении таких разложений положительно-определённые функционалы, не являющиеся характеристическими? Такие вопросы нужно решить, если пытаться описывать структуру I_0 для б. д. законов в Банаховом пространстве.

Отметим ещё приложение теоремы 1. Она может быть приложена к построению теории суммирования независимых случайных величин при отсутствии требования предельной пренебрегаемости. (см. [9], теорема 5).

§ 3. « α -разложения» безгранично делимых законов

Сформулированные аналитически, достаточные условия теоремы 1 могут годиться в более общей с аналитической точки зрения теореме (см. Ю. В. Линник, [12]).

Теорема 2. Пусть $\varphi(t)$ — х. ф. б. д. закона F с ограниченным не более, чем счётным спектром вида (2.2) и (2.3) (так что $F \in I_0$). Пусть для какой-

либо последовательности реальных значений $t_k \rightarrow 0$ имеют место соотношения:

$$(3.1) \quad (f_1(t_k))^{a_1} \dots (f_s(t_k))^{a_s} = \varphi(t_k),$$

где $a_i > 0$ ($j = 1, 2, \dots, s$); $f_j(t)$ — х. ф. случайных величин.

Тогда равенства (3.1) имеют смысл и место для всех реальных и комплексных значений t_k , и все $f_j(t)$ будут х. ф. б. д. законов со спектром входящим в спектр F .

Заметим, что если $\varphi(t)$ есть х. ф. для какого-либо закона не обязательно б. д., но такого, что $\varphi(t)$ аналитическая и не обращается в нуль в какой-либо полосе, то из равенств (3.1) вытекает, что (3.1) имеют место для любого t_k в этой полосе, и $f_j(t)$ там аналитические (см. [12], а также D. Dugué [13]).

Частный случай этой теоремы ($\varphi(t)$ х. ф. нормального закона) рассмотрен в работах [14] и [15], а также Е. LUKÁCS [16]. Он интересен тем, что из него весьма просто вытекает теорема Скитовича—DARMOIS (см. [15]). Так как эта теорема сама по себе интересна, то возникает вопрос об обобщении этого частного случая.

Пусть имеют место равенства:

$$(3.2) \quad \prod_{j=1}^{\infty} (f_j(t_k))^{a_j} = \varphi_0(t_k),$$

где $a_j > 0$ и $f_j(t)$ с. ф., а $\varphi_0(t)$ х. ф. нормального закона.

При каких условиях, налагаемых на a_j , $f_j(t)$ будут и сами х. ф. нормального закона? Можно ли усилить требование (3.2), заменив его равенством на целом сегменте $|t| \leq \delta$, но и при этом требовании решение этой задачи не удаётся. Удаётся лишь обнаружить нормальность $f_j(t)$ и верность (3.2) во всех точках t_k комплексной плоскости, если $\inf_j a_j = c > 0$. Случай

же $a_j \rightarrow 0$ не поддаётся изучению. Между тем установление нормальности $f_j(t)$ для всех j в условиях (3.2) привело бы к обобщению теоремы Скитовича—DARMOIS на линейные формы от бесконечного числа компонент:

$L_i = \sum_{j=1}^{\infty} a_{ij} X_j$ ($i = 1, 2$) (линейные функционалы от случайных процессов с

независимыми приращениями и дискретным временем). Если далее рассматривать линейные функционалы от случайных процессов с непрерывным временем, то встаёт вопрос об изучении соотношений вида:

$$(3.3) \quad \exp \int \ln f(t_{k,j}) dF(j) = \varphi_0(t_k),$$

где $\varphi_0(t)$ х. ф. нормального закона, а $f(t, j)$ х. ф. t при каждом j . Этот вопрос не разобран. Интересно также поведение $f(t, j)$, если справа стоит функция $\varphi_0(t)$ аналитическая в какой-либо полосе.

Наконец, подобные же вопросы встают для случая х. ф. со многими переменными t_1, t_2, \dots, t_s ; пусть

$$\prod_{j=1}^n (f_j(t_1^{(k)}, \dots, t_s^{(k)}))^{a_j} = \varphi_0(t_1^{(k)}, \dots, t_s^{(k)}),$$

где $\varphi_0(t_1, \dots, t_s)$ функция s комплексных переменных в полицилиндре, а $f_j(t_1, \dots, t_s)$ х. ф. и $a_j > 0$. Что нужно потребовать от последовательности $(t_1^{(k)}, \dots, t_s^{(k)}) \rightarrow (0, \dots, 0)$, чтобы обеспечить аналитичность $f_j(t_1, \dots, t_s)$?

§ 4. Об одном аналоге безграничной делимости законов

Пусть $\varphi(z) = Ee^{zx} = \int_{-\infty}^{\infty} e^{iz\xi} dF(\xi)$ х. ф. б. д. закона, и притом целая функция комплексного переменного z . В таком случае, $\varphi(z)$ не имеет нулей, так что $\varphi(z) = \exp g(z)$, где $g(z)$ целая функция. Тогда имеем: при $\xi_2 \geq \xi_1$:

$$(4.1) \quad F(\xi_2) - F(\xi_1) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{e^{-iz\xi_2} - e^{-iz\xi_1}}{-z} \varphi(z) dz \geq 0.$$

Здесь интегрирование идёт по мнимой оси.

Основным свойством б. д. законов является следующее: при любом $\alpha > 0$, образуем $(\varphi(z))^\alpha = \exp \alpha g(z)$ и заменим в интеграле (4.1) $\varphi(z)$ на $(\varphi(z))^\alpha$. Результат интегрирования будет всегда неотрицателен:

$$(4.1) \quad F_\alpha(\xi_2) - F_\alpha(\xi_1) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{e^{-z\xi_2} - e^{-z\xi_1}}{-z} (\varphi(z))^\alpha dz \geq 0.$$

Заметим, что целая функция $\varphi(z)$ будет обладать «свойством хребта»: реальная ось будет для неё «хребтом модуля», именно: $|\varphi(z)| \leq \varphi(x)$ при $z = x + iy$. Также и $(\varphi(z))^\alpha = \exp \alpha g(z)$ при $\alpha > 0$ будет обладать «свойством хребта».

Однако, «свойством хребта» будут обладать не только б. д. х. ф. функции. Например, функция:

$$(4.3) \quad \varphi(z) = \exp(z^2 + \lambda_1(e^{2z} - 1) + \lambda_2(e^{3z} - 1) - \mu(e^z - 1))$$

при заданных $\lambda_1 > 0$, $\lambda_2 > 0$ и достаточно малом $\mu > 0$ будет х. ф. (см. [12]). Однако, она, очевидно, не б. д., т. е. полагая:

$$(4.4) \quad f_\alpha(\xi) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{-z\xi} (\varphi(z))^\alpha dz, \quad \alpha > 0$$

получим: $f_\alpha(\xi) \geq 0$ при $\alpha = 1$, но это не может быть верным для всех $\alpha < 1$. Тем не менее, $(\varphi(z))^\alpha$ обладает «свойством хребта».

Оказывается, однако, «свойство хребта» при отсутствии нулей делает $\varphi(z)$ сходной с б. д. Нужно только при интегрировании выбрать в качестве контура не мнимую ось, а другой контур \mathcal{L} , связанный с методом перевала. В частности, имеет место теорема:

Теорема. Пусть $\varphi(z) = \exp g(z)$ целая чётная функция без нулей, порядка выше первого, реальная на реальной оси и имеющая «свойство хребта»:

$$|\varphi(z)| \leq \varphi(x).$$

Тогда при любом $\alpha > 0$

$$(4.5) \quad f_{\alpha}(\xi) = \frac{1}{2\pi i} \int_{\mathcal{L}} e^{-z\xi} (\varphi(z))^{\alpha} dz > 0,$$

где \mathcal{L} контур зависящий от $\varphi(z)$, но не от α и выбираемый с помощью метода наискорейшего спуска.

В случае б. д. $\varphi(z)$, контур \mathcal{L} сопоставляет х. ф. $\varphi(z)$ и $(\varphi(z))^{\alpha}$ новые вероятностные плотности, не совпадающие с прежними (которые могут и не существовать). Например, для $\varphi(z) = \exp(e^z + e^{-z} - 2)$ (4.5) даёт набор плотностей, отвечающих каждому α , а (4.2) — композицию двух законов Пуассона.

ЛИТЕРАТУРА

- [1] СКИТОВИЧ, В. П.: "Об одном свойстве нормального распределения." *ДАН СССР* **89** (1953) 217—219.
- [2] ЗИНГЕР, А. А.: "Независимость квазиполиномиальных статистик и аналитические свойства распределений." *Теория вероятностей и её применения* **3** (1958) 265—284.
- [3] Линник, Ю. В.: "О полиномиальных статистиках в связи с аналитической теорией дифференциальных уравнений." *Вестник ЛГУ* № **1** (1956) 35—48.
- [4] Линник, Ю. В.: "Линейные формы и статистические критерии I." *Украинский Мат. Журнал* **5** (1953) 207—243. II.: там же, **5** (1953) 247—290.
- [5] RÉNYI, A.: „On the algebra of distributions". *Publ. Math. Univ. Debrecen* **1** (1950) 135—149.
- [6] CRAMÉR, H.: „Problems on probability theory". *Ann. of Math. Stat.* **18** (1947) 165—193.
- [7] CRAMÉR, H.: „Über eine Eigenschaft der normalen Verteilungsfunktion". *Math. Zeitschr.* **41** (1936) 405—414.
- [8] РАЙКОВ, Д. А.: "О разложении законов Гаусса и Пуассона." *Изв. АН СССР сер. матем.* **1** (1938) 90—124.
- [9] Линник, Ю. В.: "Общие теоремы о разложении безгранично делимых законов. I." *Теория вероятностей и её применения* **3** (1958) 3—40, II.: там же, **4** (1955) 55—85, III.: там же, **4** (1959) 150—170.
- [10] CRAMÉR, H.: „On the factorization of certain probability distributions". *Arkiv. für Mat.* **1** (1949) 61—65.
- [11] ХИЛЛ, Э.: *Функциональный анализ и полугруппы*. ГГТИ, 1952.
- [12] Линник, Ю. В.: "Об « α -разложениях» безгранично делимых вероятностных законов." *Вестник ЛГУ* № **1** (1959) 14—23.
- [13] DUGUÉ, D.: „Résultats sur les fonctions absolument monotones et applications à l'arithmétique des fonctions du type positif". *C.R. Ac. Sci. Paris*, **244** (1957) 715—717.
- [14] Линник, Ю. В.: "Одна задача о характеристических функциях вероятностных распределений." *Усп. Мат. Наук* **10** (1955) 137—138.
- [15] ЗИНГЕР, А. А. — Линник, Ю. В.: "Об одном аналитическом обобщении теоремы Крамера и его применении." *Вестник ЛГУ* № **11** (1955) 51—56.
- [16] LUKÁCS, E.: „Les fonctions caracteristiques analytiques". *Annales Inst. H. Poincaré* **15** (1957) 217—251.

ЛЕКЦИЯ 5

 К ТЕОРИИ БОЛЬШИХ УКЛОНЕНИЙ ДЛЯ СУММ НЕЗАВИСИМЫХ
 СЛУЧАЙНЫХ ВЕЛИЧИН. ОДНА ЗАДАЧА ИЗ ТЕОРИИ ПРЕОБРА-
 ЗОВАНИЯ FOURIER

§ 1. Постановка задачи

Классическая теория суммирования независимых случайных величин (см. Б. В. Гнеденко, А. Н. Колмогоров [1]) как известно трактует распределения вероятностей нормированных и центрированных серий сумм. Для простейшего случая нарастающих сумм, исследованию подвергаются вероятности вида :

$$(1.1) \quad P \left\{ \frac{X_1 + X_2 + \dots + X_n}{B_n} - A_n < x \right\}$$

при соответствующем подборе констант A_n и B_n . При этом x предполагается заданным, а $n \rightarrow \infty$.

В 1938 г. Н. Срамёр [2] доказал довольно общую теорему иного типа в ней предполагалось, что x может изменяться вместе с n , например, при $n \rightarrow \infty$, x может также стремиться к ∞ . При этом можно ожидать, что (1.1) будет близко к 1, и рассматривать отношение разности 1 и (1.1) к некоторой стандартной разности. Н. Срамёр по существу применял метод «сопряжённых вероятностей», который А. Я. Хинчин [3] в (1929) ввёл для трактовки той же задачи в более узкой постановке.

В. В. Петров [4] значительно обобщил и усилил теорему Н. Срамёра. Сформулируем теорему В. В. Петрова для случая одинаково распределённых слагаемых.

Пусть Z_1, Z_2, \dots независимые одинаково распределённые случайные величины с интегральным законом распределения $V(y)$. Пусть выполняется условие Н. Срамёра : интеграл

$$(1.2) \quad R(h) = \int_{-\infty}^{\infty} e^{hy} dV(y) < \infty \quad \text{при} \quad |h| \leq A.$$

Пусть $E Z_i = 0$; $D(Z_i) = \sigma^2$; составим нормированную сумму $S_n = \frac{Z_1 + Z_2 + \dots + Z_n}{\sigma \sqrt{n}}$ с законом распределения $F_v(x)$.

Теорема В. В. Петрова. Пусть $1 < |x| < \eta_0 \sqrt{n}$, где η_0 достаточно малая константа. Тогда имеем :

$$(1.3) \quad \frac{1 - F_n(x)}{1 - \Phi(x)} = \exp \left[\frac{x^3}{\sqrt{n}} \lambda \left(\frac{x}{\sqrt{n}} \right) \right] (1 + B_1 \eta_0), \quad \text{если} \quad x \rightarrow +\infty,$$

$$(1.4) \quad \frac{F_n(x)}{\Phi(x)} = \exp \left[-\frac{x^3}{\sqrt{n}} \lambda \left(-\frac{x}{\sqrt{n}} \right) \right] (1 + B_2 \eta_0), \quad \text{если} \quad x \rightarrow -\infty,$$

здесь $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$; B_1, B_2 — ограниченные числа; $\lambda(z) = \sum_{v=0}^{\infty} a_v z^v$ — степенной ряд, сходящийся в круге $|z| < \rho_0$.

Коэффициент a_v выражается через первые v семиинвариантов (и стало быть, моментов) распределения $V(y)$ (по поводу приложения этой теоремы в арифметике кватернионов см. [5]).

Рассмотрим постановку задачи в простейшем случае центральной предельной теоремы и случае одинаково распределённых слагаемых с нулевым средним. Если такие слагаемые имеют дисперсию σ^2 , то нормированные суммы

$$\frac{X_1 + \dots + X_n}{\sqrt{n}}$$

будут притягиваться к одному и тому же нормальному закону $\mathcal{N}(0, \sigma)$; он как бы собирает в своей области притяжения все случайные величины с нулевым средним и заданной дисперсией σ^2 .

Будем требовать такого феномена как бы собирания в одну область притяжения всех распределений $V(y)$ с заданными первыми моментами $0, \mu_1, \mu_2, \dots, \mu_k$. Как видно из формул (1.3) и (1.4), такое «собираение» при выполнении условия Скрамёра (1.2) будет происходить, вообще говоря, только при условии: (считая $x > 0$):

$$(1.5) \quad x < n^{1/2-\varepsilon}, \quad \varepsilon > \frac{1}{k+4},$$

ибо только тогда $\frac{x^3}{\sqrt{n}} \left(\frac{x}{\sqrt{n}} \right)^{k+1} \rightarrow 0$ при $n \rightarrow \infty$.

Это приводит нас к тому, что при попытках построения общей теории больших уклонений, для получения феномена «собираения в области притяжения», надлежит рассматривать сегмент $|x| < n^{1/2-\varepsilon}$ при $\varepsilon > 0$; при больших сегментах, постановка задач о предельных теоремах для больших уклонений должна как-то отличаться от рассмотрения «областей притяжения» аналогичных классической теории.

§ 2. Результаты W. Richter-a

W. РИХТЕР [6], (1957), систематически исследовал большие уклонения методом перевала. Оказалось, что все предыдущие исследователи больших уклонений по существу, в неявной форме применяли метод перевала. W. РИХТЕР впервые доказал локальные теоремы для больших уклонений в условии Н. СКАРМЭРА (1.2), в случае одинаковых и неодинаковых слагаемых, непрерывных и решётчатых.

Он обнаружил, в известном смысле, необходимость условия Н. СКАРМЭРА для существования формул вида (1.3) и (1.4).

Подробные формулировки теорем W. РИХТЕР-а находятся в указанных работах. Дадим краткую схему применяемого им метода, что будет существенным для дальнейшего. Выводится локальная теорема с большими уклонениями, для непрерывно распределённых слагаемых. Для простоты,

предположим, что $V'(x) = g(x)$ существует всюду и ограничена (у W. Рис-тер-а менее жёсткие условия). Положим :

$$M(z) = \int_{-\infty}^{\infty} e^{zx} g(x) dx; \quad z = u + iy.$$

В силу условия Н. СРАМЁР-а (1.2), $M(z)$ регулярна в полосе $|x| < A$. Пусть $p_n(x)$ плотность вероятности для нормированной суммы, тогда :

$$(2.1) \quad p_n(x) = \frac{\sigma \sqrt{n}}{2\pi i} \int_{c-i\infty}^{c+i\infty} (M(z))^n e^{-\sigma \sqrt{n}zx} dz.$$

Далее вступает в действие аналитичность $M(z)$, обеспеченная условием Н. СРАМЁР-а (1.2). Контур $(c - i\infty, c + i\infty)$ переносится в точку перевала, которая будет близка к 0 вместе с $\frac{x}{\sqrt{n}}$, и после этого (2.1) рассчитывается по методу перевала.

Мы видим, что здесь существенна аналитичность $M(z)$, т. е. условие Н. СРАМЁР-а (1.2).

§ 3. Возможный подход при невыполнении условия Н. СРАМЁР-а. Одна задача об интеграле Fourier

При невыполнении условия Н. СРАМЁР-а нет аналитичности $M(z)$, и предыдущий метод неприменим. Мы имеем всё же :

$$(3.1) \quad p_n(x) = \frac{\sigma \sqrt{n}}{2\pi i} \int_{-i\infty}^{i\infty} (M(z))^n \exp(-\sigma \sqrt{n}zx) dz.$$

Хотя $M(z)$ и не аналитическая, может случиться, что $M(z)$ при $y \geq 0$ и при $y < 0$ продолжима в плоскость комплексного переменного, но по-разному, т. е. разные лучи : $y \geq 0$ и $y < 0$ дают разные продолжения.

Простейшим примером является плотность Cauchy :

$$g(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}; \quad M(z) = e^{-|z|} \text{ при } z = iy.$$

Значит, при $y \geq 0$, $M(z) = e^{iz}$; при $y < 0$, $M(z) = e^{-iz}$. Функции e^{iz} и e^{-iz} целые функции.

Другой пример :

$$g(x) = \frac{A_0}{(1+x^2)(4+x^2)}. \quad \text{Здесь существует дисперсия, и } M(z) = 2e^{-|z|} - e^{-2|z|} \text{ (} z = iy \text{)}.$$

Продолжение с луча $y \geq 0$ даёт целую функцию $2e^{iz} - e^{2iz}$; продолжение с луча $y < 0$ — целую функцию $2e^{-iz} - e^{-2iz}$. Этот же эффект — продолжение с двух лучей $y \geq 0$ и $y < 0$ с помощью двух различных целых

функций будет и при любой рациональной плотности вероятности $r(x)$. Подобное обстоятельство позволяет сделать и расчёт вероятности больших уклонений. Вертикальный контур $x = 0$ надо заменить *изломанным*: луч $y \geq 0$ заменяется отрезком оси абсцисс: $0 \leq x \leq u_0$ и вертикалью: $z = u_0 + iy, y \geq 0$; луч $y < 0$ заменяется отрезком $-u_0 \leq u \leq 0$ и вертикалью $z = -u_0 + iy, y < 0$.

При этом можно снова выделить точки перевала (при $\frac{x}{\sqrt{n}} \rightarrow 0$ они весьма близки к 0 и ещё ближе к оси абсцисс), и рассчитать интегралы по методу перевала. Любопытно, что при этом, роль моментов при условии Н. Стаммэ-а играют k -е произведения в 0 двух различных целых продолжений. Их можно назвать *ложными моментами*. Если $x \leq n^{\frac{1}{2}-\varepsilon}$, происходит феномен собирания плотностей $g(y)$ в область притяжения при данных *ложных моментах*.

Заметим ещё, что если $\Delta(x)$ непрерывная функция под условиями:

$$(3.2) \quad \Delta(x) = O(e^{-\alpha_0|x|}) \text{ при } |x| \rightarrow \infty; \quad \int_{-\infty}^{\infty} \Delta(x) dx = 0,$$

и $g(x) > |\Delta(x)|$, то $g(x) + \Delta(x)$ будет вероятностной плотностью, и $\int_{-\infty}^{\infty} e^{zx} \Delta(x) dx$ будет продолжимым в полосу $|u| < \alpha_0$ ($z = u + iy$). Ввиду этого, наряду с рациональными плотностями $r(x)$ можно рассматривать плотности вида:

$$(3.3) \quad r(x) + \Delta(x)$$

и можно проводить указанные выше рассуждения. Таким образом, можно получить в частности, теоремы о больших уклонениях для рациональных плотностей, сколь угодно «испорченных» на любом конечном интервале любыми искажениями.

Предыдущее наводит на мысль о рассмотрении «опорных» плотностей, которые вели себя подобно рациональным. Это приводит к задаче:

Пусть $g(x) \in L_1$. Рассматривается преобразование Fourier:

$$(3.4) \quad \int_{-\infty}^{\infty} e^{iyx} g(x) dx = \varphi(y).$$

Что нужно потребовать от $g(x)$ для того, чтобы с двух лучей: $y \geq 0$ и $y < 0$ существовали (возможно различные) аналитические продолжения $\varphi(y)$, которые охватывали бы окрестность 0 и имели бы там лишь конечное число изолированных особенностей.

В частности, когда эти продолжения будут целыми функциями? Если $g(x) \in L_1$ вероятностная плотность, допускающая такие продолжения $\varphi(y)$, то можно получать теоремы о больших уклонениях предыдущим способом. Можно сколь угодно «испортить» её добавлением описанной выше $\Delta(x)$. Предельные теоремы будут формулироваться в терминах «ложных моментов».

Если $|x| \leq n^{1/2-\varepsilon}$, происходит феномен «собирания» в области притяжения по количеству $\sim \frac{1}{\varepsilon}$ первых «ложных моментов».

«Опорных» плотностей этого типа существует много. Все алгебраические вероятностные плотности, все плотности вида $Ae^{-|x-\beta|^\alpha}$ и их линейные комбинации будут «опорными». Желательно, однако, расширить их класс, исследуя поставленную выше задачу об интеграле Fourier.

Тогда можно пытаться подойти к задаче построения общей теории больших уклонений для одинаковых слагаемых следующим образом.

В сумме $X_1 + \dots + X_n$ маловероятны очень большие значения многих слагаемых; подавляющее большинство слагаемых будет лежать не в очень большом интервале. Там нужно пытаться аппроксимировать плотность с точностью до $\Delta(x) = O(e^{-a_0|x|})$ набором «опорных» плотностей, и после этого делать расчёт соответствующих интегралов описанным выше методом.

Предельная теорема должна формулироваться в терминах ложных моментов для «опорных» плотностей; «собирание в области притяжения» будет при $|x| \leq n^{1/2-\varepsilon}$.

Мы разбирали здесь локальные теоремы; интегральные теоремы должны получаться при менее жестких условиях.

ЛИТЕРАТУРА

- [1] ГНЕДЕНКО, Б. В. — КОЛМОГОРОВ, А. Н.: *Предельные распределения для сумм независимых случайных величин*. ГТТИ, 1949.
- [2] СРАМЕР, Н.: "Sur un nouveau théorème-limite de la théorie des probabilités". *Actual. Sci. et ind., Paris*, 736 (1938).
- [3] ХИНЧИН, А. Я.: *Математические основания статистической механики* М.—Л. 1943.
- [4] ПЕТРОВ, В. В.: "Обобщение предельной теоремы Крамера." *УМН*. 9 (1954) 195—202.
- [5] ЛИННИК, Ю. В.: "Применение теории цепей Маркова в арифметике кватернионов." *УМН* 9 (1954) 203—210.
- [6] РИХТЕР, В.: "Локальные предельные теоремы для больших уклонений." *Теория вероятностей и её применения* 2 (1957) 214—229.

ÖT ELŐADÁS A VALÓSZÍNŰSÉGSZÁMÍTÁS ÉS A SZÁMELMÉLET BIZONYOS PROBLÉMÁIRÓL

JU. V. LINNIK

1. ELŐADÁS

A SZÓRÁS MÓDSZER ALKALMAZÁSA BIZONYOS BINÁRIS ADDITÍV PROBLÉMÁK MEGOLDÁSÁRA

A szerző $n = \varphi + D'v$ típusú bináris additív problémák egy aszimptotikus megoldását dolgozza ki aránylag gyenge feltételek mellett, egy új módszer segítségével. Itt $\{\varphi\}$ a természetes számok egy sorozata, D' és v a természetes számok egy rendszerén fut végig. Bevezet egy TURÁN típusú szórását és felhasználja VINOGRADOV módszerének elemeit, valamint a CSEBISEV egyenlőtlenséget. Egy alapvető példa az az eset amikor $\varphi = \varphi(x, y)$ egy kvadrátikus forma.

2. ELŐADÁS

**A SZÓRÁS MÓDSZER ALKALMAZÁSAI. OSZTÓK PROBLÉMÁJA.
A HARDY—LITTLEWOOD PROBLÉMA. MÁS PROBLÉMÁK**

A szerző a szórás módszert alkalmazza bizonyos bináris problémák megoldására. Ezek tartalmazzák az osztók problémáját: $\sum_{n \leq x} \tau_2(n) \tau_k(n+l)$ ($k=2,3,\dots$, aszimptotikus becslését. A szórás módszert ezenkívül az $n = \xi^2 + \eta^2 + p_1 p_2^a$ és az $n = \xi^2 + \eta^2 + p_1 d^{m^a}$ (d fix $a \geq 1$ és fix p_1, p_2 primek, $p_1 \leq n^{1-a}$, $p_2^a \leq n^a$, $d^{m^a} \leq n^a < 1/2$), továbbá az $n = \xi^2 + \eta^2 + p$ (p prím) (Hardy—Littlewood egyenlet) egyenletek és más problémák megoldására használja. Az első két egyenlet megoldása a Riemann-féle sejtés segítségével sem volt ismeretes.

3. ELŐADÁS

**GÖMB RÁCSPONTJAI ÉS A MARKOV LÁNCOK. ERGODIKUS TÉTELEK
EGÉSZ MÁTRIXOKRA VONATKOZÓ ANALOGONJAI**

Egy gömb rácsponjtjainak ergodikusságával foglalkozik. Rámutat, hogy ezek okozzák a gömb ezen pontjainak aszimptotikusan egyenletes eloszlását. Ezen tulajdonságok általánosíthatók az $\frac{n^2 + n}{2}$ dimenziós tér n -ed fokú felületeinek egy osztályára és ezzel bizonyos kapcsolatokat nyerünk az algebrai számelmélet, az egész mátrixok és néhány Markov láncokra vonatkozó határeloszlástétel között.

4. ELŐADÁS

KORLÁTLANUL OSZTHATÓ ELOSZLÁSOK NÉHÁNY TULAJDONSÁGA

Néhány „exact” (nem aszimptotikus) valószínűség-számítási problémával foglalkozik. Nevezetesen a következő problémával: mi a feltétele annak, hogy két lineáris statisztika L_1 és L_2 egyforma normális eloszlású és ekvivalens legyen (MARCINKIEWICZ—RÉNYI típusú probléma). Ezenkívül CRAMÉR korlátlanul osztható eloszlások felbontására vonatkozó problémájával foglalkozik, valamint ennek egy analitikus irányba történő általánosításával („ α -felbontás”). Banach térből vett értékű valószínűségi változók karakterisztikus funkcionáljaira vonatkozó néhány problémát vet fel és a korlátlanul osztható eloszlások egy analogonját adja meg.

5. ELŐADÁS

**MEGJEGYZÉSEK FÜGGETLEN VALÓSZÍNŰSÉGI VÁLTOZÓK ÖSSZEGEI
NAGY INGADOZÁSAINAK ELMÉLETÉHEZ; EGY PROBLÉMA A FOURIER
TRANSZFORMÁCIÓVAL KAPCSOLATBAN**

A nagy ingadozások elméletének néhány eredményét ismerteti és néhány gondolatot, amelynek segítségével egy általános elmélet dolgozható ki. Felvet egy a Fourier transzformáltak bizonyos félsíkokon való analitikus folytatására vonatkozó problémát.

FIVE LECTURES ON SOME TOPICS IN NUMBER THEORY AND PROBABILITY THEORY

Ю. В. LINNIK

Abstracts

LECTURE 1

THE VARIANCE METHOD FOR SOLVING CERTAIN BINARY ADDITIVE PROBLEMS

To give an asymptotic solution of a binary additive problem of the type $n = \varphi + D'v$ where $\{\varphi\}$ is a sequence of natural numbers and D', v run over systems of natural numbers under comparatively mild conditions, a new method is set forth. This introduces a variance of P. TURÁN's type and uses the elements of I. M. VINOGRADOV's method and TSHEBYCHEV's inequality. The case of $\varphi = \varphi(x, y)$ being a binary quadratic form is one of the principal examples.

LECTURE 2

APPLICATIONS OF THE VARIANCE METHOD. THE DIVISOR PROBLEM THE HARDY—LITTLEWOOD PROBLEM. OTHER PROBLEMS

The application of the variance method to the solution of certain binary problems is given.

These include the divisor problem: the asymptotic evaluation of $\sum_{n \leq x} \tau_2(n) \tau_k(n+l)$ ($k = 2, 3, \dots$). The variance method is applicable to the study of the equations of the type $n = \xi^2 + \eta^2 + p_1 p_2^a$ and $n = \xi^2 + \eta^2 + p_1 d^{m^a}$ (d being fixed, $a \geq 1$ any fixed numbers, p_1 and p_2 primes, $p_1 \leq n^{1-a}$, $p_2^a \leq n^a$, $d^{m^a} \leq n^a d < 1/2$) the HARDY—LITTLEWOOD equation $n = \xi^2 + \eta^2 + p$ (p prime) and some other problems. The solvability of the first two equations cannot be deduced from Riemann's hypothesis at present.

LECTURE 3

INTEGER POINTS ON A SPHERE AND MARKOV CHAINS. ANALOGON OF THE ERGODIC THEOREMS FOR INTEGER MATRICES

The ergodic properties of integer points on a sphere are set forth. These are seen to be the cause of the asymptotically uniform distribution of these points on the sphere. These properties can be generalized for a class of surfaces of degree n in an $\frac{n^2 + n}{2}$ -dimensional space and represent some connection between algebraic number theory, integer matrices and some limit theorem for the Markov-chain theory.

LECTURE 4

SOME PROPERTIES OF INFINITELY DIVISIBLE LAWS

Some "exact" (not asymptotic) probabilistic problems are considered. These include the problem of finding the conditions for two linear statistics L_1 and L_2 such that the equidistribution of L_1 and L_2 should be equivalent to normality of observations (problems of MARCINKIEWICZ—RÉNYI type). Cramér's problem on the decomposition of infinitely divisible laws is then considered, together with its generalization in an analytical direction (" α -decompositions"). Some problems concerning characteristic functionals for Banach-space valued random variables are proposed. An analogon of infinitely divisible laws is set forth.

LECTURE 5

CONTRIBUTIONS TO THE THEORY OF BIG DEVIATIONS FOR SUMS OF INDEPENDENT RANDOM VARIABLES. A PROBLEM ON FOURIER-TRANSFORMS

Some results on the theory of big deviations are stated, and some directions along which a general theory might be developed are discussed. A problem on the analytical continuation of Fourier-transforms from certain half-lines is proposed.

AN ASSOCIATIVITY THEOREM FOR ALTERNATIVE RINGS

by

GYÖRGY GRÄTZER and E. TAMÁS SCHMIDT

It is well-known that (with the aid of the usual method due to HILBERT) every projective plane satisfying the special Desargues' theorem may be co-ordinated with an alternative division ring (of characteristic not 2). So it occurred in the geometry the necessity of the investigation of a non-associative ring.¹

The ring R is called alternative if for all $\alpha, \beta \in R$

$$(\alpha\alpha)\beta = \alpha(\alpha\beta) \quad \text{and} \quad (\beta\alpha)\alpha = \beta(\alpha\alpha).$$

The reason why the name alternative has been given is that the associator function

$$(\alpha, \beta, \gamma) = (\alpha\beta)\gamma - \alpha(\beta\gamma)$$

is an alternative function of its variables,² that is,

$$(\alpha, \beta, \gamma) = -(\alpha, \gamma, \beta) = -(\beta, \alpha, \gamma) = -(\gamma, \beta, \alpha).$$

The co-ordinate ring R is associative if and only if Desargues' theorem holds upon the projective plane. Therefore the consideration of associative subrings of alternative rings is of importance. At first E. ARTIN proved a theorem of this kind: in an alternative ring any subring generated by two elements is associative. Later on M. ZORN [5] generalized Artin's theorem. Further generalization was achieved by R. MOUFANG [2] for some special type of alternative rings, and this result was extended to the general case by M. F. SMILEY [4]. All the proofs need many computations. R. H. BRUCK and E. KLEINFELD [1] established a sufficient condition for the associativity of a subring which is a far-reaching generalization of the earlier results. The proof of this theorem — though it is not simple — is essentially simpler than that of the earlier ones. The proof of the theorem of BRUCK and KLEINFELD is simplified in nuances in the book [3] of G. PICKERT.

Our aim is to get a necessary and sufficient condition (the former ones were only sufficient) which contains all the conditions so far mentioned as special cases. The proof of this theorem is comparatively simple and requires

¹ Non-associative ring means a not necessarily associative ring. From now on we use the term "ring" for "non-associative ring".

² In rings of characteristic not 2 this property used to serve as a definition of alternative rings.

almost no computation. From this theorem we shall derive a further theorem of BRUCK and KLEINFELD [1] (which was not to be deduced from their associativity condition as a consequence of its being only sufficient).

If A , B and C are subsets of the alternative ring R , (A, B, C) denotes the set of all associators (a, β, γ) with $a \in A$, $\beta \in B$, $\gamma \in C$. Let $D = \bigvee_{i=1}^n A_i$, $n \geq 2$, the union of the subsets A_i of R . Let A_{ij} ($i \neq j$, i and $j \leq n$, $n > 2$) consist of products of at most $n - 2$ factors taken from different A_k -s, $k \neq i$, $k \neq j$. If $n = 2$ or $i = j$ we put $A_{ij} = 0$. Finally we define \bar{D} as the subring of R generated by D . Now, the general associativity condition is as follows:

Theorem. Let R be an alternative ring and $D = \bigvee_{i=1}^n A_i$ ($n \geq 2$) a subset of R . D generates an associative subring \bar{D} of R if and only if the following two conditions are fulfilled:

- (1) $(A_i, A_i, \bar{D}) = 0$, (i and $j = 1, 2, \dots, n$)
 (2) $(A_i, A_j, A_{ij}) = 0$.

The **proof** is based on the following observation:

(*) If $(a_1, \beta, \gamma) = (a_2, \beta, \gamma) = (a_1, a_2, \beta) = 0$

then

$$(a_1 a_2, \beta, \gamma) = (a_1, \beta, a_2 \gamma) = - (a_1, \beta, \gamma a_2).$$

For the proof of (*) we define — following [1] — a function as follows

$$f(w, x, y, z) = (wx, y, z) - (x, y, z)w - x(w, y, z).$$

It is easy to verify that this function is multi-linear and skew-symmetric in its variables (i.e. it changes sign when two of its neighbouring variables are interchanged). Using these facts and the hypotheses we get

$$(a_1 a_2, \beta, \gamma) = a_2(a_1, \beta, \gamma) + (a_2, \beta, \gamma)a_1 + f(a_1, a_2, \beta, \gamma) = f(a_1, a_2, \beta, \gamma).$$

On the other hand, in a similar way

$$(a_1, \beta, a_2 \gamma) = f(a_2, \gamma, \beta, a_1).$$

Since $f(w, x, y, y)$ is skew-symmetric, we have $f(a_1, a_2, \beta, \gamma) = -f(\gamma, a_2, \beta, a_1) = -f(a_2, \gamma, \beta, a_1)$. Comparing the three equalities we get the first part of (*). Similarly we get the second part of (*).

Now, we turn to the proof of our theorem. The necessity of the conditions (1) and (2) is obvious, for the associativity of \bar{D} means — according to the definition of the associator — that $(\bar{D}, \bar{D}, \bar{D}) = 0$ and (1) and (2) are trivial consequences.

On the other hand, let us suppose the validity of (1) and (2); then we prove the associativity of \bar{D} . Let D^k denote the set of all products of k factors from D . Obviously, \bar{D} is the module generated by all the D^k ($k = 1, 2, \dots$). Since the associator function satisfies

$$(A_1 + A_2, B, C) = (A_1, B, C) + (A_2, B, C)$$

it is enough to prove that the associator function vanishes on the elements of $\bigvee_{k=1}^{\infty} D^k$. Let u, v, w be positive integers; we use an induction on $u + v + w$ to verify $(D^u, D^v, D^w) = 0$. If $u + v + w = 3$ (which is the smaller possible value of $u + v + w$) we have to prove $(D, D, D) = 0$. This is trivial because every element of (D, D, D) is in the left side of (1) or (2), and so it is equal to 0. Now, we suppose the assertion is proved for $u + v + w < N$; let $u + v + w = N$ and $\alpha \in D^u, \beta \in D^v, \gamma \in D^w$. The elements α, β and γ are products of N factors altogether. Suppose, two of these factors are from the same A_i , say from A_1 . Then we apply (*) repeatedly. From the induction hypothesis it follows that the assumptions of (*) are fulfilled. We decompose always that element³ (or those elements) which contains a factor from A_1 ; and that factor is put in the first, later in the second argument which contains a factor from A_1 . All other factors are gathered up in the third argument. Proceeding in this way we get finally a position, where in the first and in the second argument of the associator there will be elements of A_1 and so by (1) in this case the associator is 0.

Now let us suppose that all the N factors of α, β and γ are from different A_i -s. Then applying repeatedly (*) again, we let in the first and in the second arguments only single factors and all the others are gathered up in the third argument. Then by the hypothesis if in the first argument the factor is from A_i , whilst that in the second from A_j , then in the third argument the element belongs to A_{ij} , and so the associator is 0 according to (2). This completes the proof of the theorem.

We mention that the theorem might be generalized to the case when the number of the A_i -s is infinite. Naturally, then there is no bound on the number of factors of the elements of A_{ij} . Otherwise the theorem remains true, without any alteration.

Next, we turn our attention to special cases of this theorem.

Corollary 1. (Theorem of BRUCK and KLEINFELD.) *Let A_1, A_2, A_3 be subsets of an alternative ring R such that $(A_1, A_1, R) = (A_2, A_2, R) = (A_3, A_3, R) = (A_1, A_2, A_3) = 0$. Then the subset $D = A_1 \cup A_2 \cup A_3$ is contained in an associative subring of R .*

In case $A_1 = A_2$ this assertion was first proved by M. ZORN.

Indeed in case $n = 3$ A_{12} means A_3 , thus the hypotheses of our theorem are fulfilled and we get the theorem of BRUCK and KLEINFELD.

Corollary 2. *Let D be a subset of an alternative ring R and a_v ($v \in \Omega$) the elements of D . If $D_{\mu\omega}$ ($\mu, \omega \in \Omega$) denotes the set of the products of the a_v -s with a_μ and a_ω excluded, then D generates an associative subring in R if and only if $(a_\mu, a_\omega, D_{\mu\omega}) = 0$ for all $\mu, \omega \in \Omega, \mu \neq \omega$.*

A special case of Corollary 2 is

Corollary 3. (Theorem of SMILEY.) *The elements α, β, γ of the alternative ring R generate an associative subring if and only if $(\alpha, \beta, \gamma) = 0$.*

³ An element of R may be considered as a product of one factor; multiplying a product of k by one of l factors, we get a product of $k + l$ factors. So a product of n factors always might be decomposed to two elements, both of them having less than n factors.

By the aid of our theorem we can easily prove also the following assertion of BRUCK and KLEINFELD [1]:

Corollary 4. *Let A be an associative subring of the alternative ring R , and B a subset of R subject to $(A, A, B) = (B, B, R) = 0$. Then $D = A \cup B$ generates an associative subring of R .*

Proof. In case $n = 2$ condition (2) is trivially satisfied, further $(B, B, R) = 0$ is now supposed, hence in order to prove the validity of the conditions of the theorem it is enough to see that $(A, A, \bar{D}) = 0$. Let $\gamma \in \bigvee_{k=1}^{\infty} D^k$ and denote

by n the least integer such that γ is the product of n factors taken from A and B . We choose $a_1, a_2 \in A$ and by an induction on n we prove $(\gamma, a_1, a_2) = 0$ which will complete the proof. In case $n = 1$, γ is from A or from B and so $(\gamma, a_1, a_2) = 0$ follows from the associativity of A , resp. from $(A, A, B) = 0$. If the assertion is true for all $k < n$, then we again apply (*): decomposing γ we can modify the associator (γ, a_1, a_2) in such a way, that all the factors of γ which are taken from B are gathered up in the third argument, while the others are collected in the first argument. Since the associator (without altering its value) will be an element of (A, A, B) , it is equal to 0.

We note — following [1] — that Corollary 4 immediately implies the following result:

All the maximal associative subrings of the alternative ring R are maximal associative subsets, that is, they can not be enlarged so that all the associators should remain 0.

Indeed, in case the maximal associative subring A can be enlarged with the element β so that the associators remain 0, then $(A, A, \beta) = 0$ and obviously $(\beta, \beta, R) = 0$, so, by Corollary 4, A and β generate an associative subring, in contradiction to the maximality of A .

(Received April 7, 1959.)

REFERENCES

- [1] BRUCK, R. H. and KLEINFELD, E.: "The structure of alternative division rings" *Proc. Amer. Math. Soc.* **2** (1951) 878—890.
- [2] MOUFANG, R.: "Zur Struktur der projektiven Geometrie der Ebene". *Math. Ann.* **105** (1931) 536—601.
- [3] PICKERT, G.: *Projektive Ebene*. Springer Verlag, 1955.
- [4] SMILEY, M. F.: "Some questions concerning alternative rings". *Bull. Amer. Math. Soc.* **57** (1951) 36—43.
- [5] ZORN, M.: "Theorie der alternativen Ringe". *Abh. Math. Sem. Hamburgischen Univ.* **8** (1930) 123—147.

EGY ASSZOCIATIVITÁSI TÉTEL ALTERNATÍV GYŰRŰKRE

GRÄTZER GY. és SCHMIDT E. T.

Kivonat

Legyen R egy alternatív gyűrű. Ha a, β, γ elemei R -nek, akkor (a, β, γ) jelöli az $(a\beta)\gamma - a(\beta\gamma)$ elemet. Az (a, β, γ) függvényt asszociátor-függvénynek nevezzük, s R alternativitása azt jelenti, hogy az asszociátor-függvény változóinak alternatív függvénye. Ha A, B, C részhalmazai R -nek, akkor (A, B, C)

jelölje az összes olyan (α, β, γ) asszociátor halmazát, amelyre $\alpha \in A$, $\beta \in B$, $\gamma \in C$. Legyenek az A_i -k az R gyűrűk részhalmazai, s $D = \bigvee_{i=1}^n A_i$ ($n \geq 2$).

Definiáljuk az A_{ij} részhalmazt, hogy $i \neq j$, $n > 2$ esetén álljon mindazon legfeljebb $n - 2$ tényezős szorzatokból, amelyeknek minden tényezőjét különböző A_i -kből vettük, s $i = j$ vagy $n = 2$ esetén legyen $A_{ij} = 0$. Végezetül legyen \bar{D} az R gyűrűnek a D részhalmaz által generált részgyűrűje. A dolgozat célja, hogy bebizonyítsa a következő állítást:

Tétel. *Legyen R egy alternatív gyűrű, A_i részhalmazai R -nek, s $D = \bigvee_{i=1}^n A_i$ ($n \geq 2$). D akkor és csak akkor generálja R -nek asszociatív részgyűrűjét, ha a következő két feltétel teljesül:*

- (1) $(A_i, A_i, \bar{D}) = 0$, ($i, j = 1, 2, \dots, n$)
- (2) $(A_i, A_j, A_{ij}) = 0$.

A tétel bizonyítása a következő lemmából egyszerűen adódik:

Tegyük fel, hogy az R alternatív gyűrű a_1, a_2, β, γ elemeire

$$(a_1, \beta, \gamma) = (a_2, \beta, \gamma) = (a_1, a_2, \beta) = 0,$$

akkor

$$(a_1 a_2, \beta, \gamma) = (a_1, \beta, a_2 \gamma) = - (a_1, \beta, \gamma a_2).$$

A bizonyítás egyszerű teljes indukciós megfontolással, minden számolás nélkül adódik.

A tételből az $n = 3$ esetben nyerhető BRUCK és KLEINFELD [1] asszociativitás-feltétele, s tovább specializálva SMILEY és ZORN tételei. Szintén egyszerűen nyerhető BRUCK és KLEINFELD [1] egy további tétele is. A dolgozatban szereplő tétel bizonyítása egyszerűbb, mint az eddigi speciális eseteké.

ОДНА ТЕОРЕМА ОБ АССОЦИАТИВНОСТИ ДЛЯ АЛЬТЕРНАТИВНЫХ КОЛЕЦ

GY. GRÄTZER и E. T. SCHMIDT

Резюме

Пусть R есть альтернативное кольцо. Если α, β, γ элементы R , то (α, β, γ) обозначает элемент $(\alpha\beta)\gamma - \alpha(\beta\gamma)$. Функция (α, β, γ) называется ассоциатор-функцией, и альтернативность R означает альтернативность этой функции. Если A, B, C подмножества R , то пусть (A, B, C) обозначает множество всех ассоциаторов (α, β, γ) , для которых $\alpha \in A, \beta \in B, \gamma \in C$.

Пусть A_i суть подмножества R и $D = \bigvee_{i=1}^n A_i$ ($n \geq 2$). Определим подмножество A_{ij} , чтобы при $i \neq j$, $n > 2$ оно состояло из всех произведений $n - 2$ сомножителей, каждый из которых принадлежит различным A_i , а в случае $i = j$ или $n = 2$ $A_{ij} = 0$. Наконец, пусть \bar{D} будет подкольцо кольца R , генерированное множеством D . Цель работы доказать следующее утверждение:

Теорема. Пусть R альтернативное кольцо, A_i его подмножества, $D = \bigvee_{i=1}^n A_i$ ($n \geq 2$). D в том и только в том случае генерирует ассоциативное подмножество R , если выполняются следующие два условия:

- (1) $(A_i, A_i, \overline{D}) = 0$, ($i, j = 1, 2, \dots, n$)
 (2) $(A_i, A_j, A_{ij}) = 0$.

Доказательство теоремы легко получается из следующей леммы: Предположим, что для элементов $\alpha_1, \alpha_2, \beta, \gamma$ альтернативного кольца R

$$(\alpha_1, \beta, \gamma) = (\alpha_2, \beta, \gamma) = (\alpha_1, \alpha_2, \beta) = 0,$$

тогда

$$(\alpha_1 \alpha_2, \beta, \gamma) = (\alpha_1, \beta, \alpha_2 \gamma) = -(\alpha_1, \beta, \gamma \alpha_2).$$

Доказательство получается простым соображением с помощью математической индукции без всякого счёта.

Из теоремы в случае $n = 3$ может быть получено условие ассоциативности BRUCK-а и KLEINFELD-а [1] и дальнейшей специализацией теорема SMILEY и ZORN-а. Легко получить и ещё одну теорему BRUCK-а и KLEINFELD-а [1]. Доказательство теоремы проще, чем доказательство этих специальных случаев.

LINEÁRIS EGYENLETRENDSZEREK MEGOLDÁSA SZIMPLEX MÓDSZERREL

Bevezetés

KREKÓ BÉLA

A gazdasági életben felmerülő matematikai problémák megoldása számos esetben a *lineáris programozás* néven ismert módszerre vezethető vissza [1]. A lineáris programozásnál egy olyan elsőfokú függvény extrémális értékének meghatározásáról van szó, amelynek értelmezési tartományát egy — az adott gazdasági feltételeket reprezentáló — lineáris egyenlőtlenségrendszer szabja meg. Ennek következtében a kérdéses értelmezési tartomány olyan konvex halmazt alkot, amelynek véges számú csúcspontja van. Ha a vizsgálandó függvénynek van extrémális értéke, azt e csúcspontok valamelyikében feltétlenül felveszi. A megoldásra szolgáló módszerek közül legelterjedtebb a G. B. DANTZIG-tól származó *simplex módszer* [2]. A simplex módszernek az az alapgondolata, hogy az értelmezési tartomány egyik csúcspontjából kiindulva úgy haladjunk csúcsponttól csúcspontra, hogy közben a függvény helyettesítési értékei monoton sorozatot alkossanak. S mivel a csúcspontok száma véges, a kérdéses extrémum — amennyiben létezik — véges számú lépésben meghatározható.

A simplex módszer elemzése közben kiderült, hogy annak alapvető mozzanata az az eljárás, amelyet a továbbiakban *elemi transzformációnak* fogunk nevezni. Ez olyan speciális lineáris transzformációt jelent, amelynek révén egy adott bázisból olyan új bázisba térünk át, amely csupán egyik vektorában különbözik az eredetitől. S mivel bármely lineáris transzformáció elemi transzformációk sorozatára bontható fel, érthető, hogy ezen az úton lehetőség nyílik számos lineáris algebrai probléma egyszerű — determinánsmentes — tárgyalására.¹

Ebben a dolgozatban csupán a lineáris egyenletrendszerek megoldásával kívánunk foglalkozni. A bemutatandó eljárással kapcsolatban, amelyet az előzőek alapján ugyancsak *simplex módszernek* fogunk nevezni, előrebocsátjuk a következőket:

a) Közvetlenül alkalmazható bármilyen végesdimenziójú lineáris egyenletrendszer megoldására. (Az inhomogén rendszereket nem kell előbb átalakítani homogén rendszerekké.)

b) A kompatibilitás és a rang kérdésének eldöntésével egyidejűleg explicit formában kapjuk meg a megoldást is. Nincs tehát szükség további rekurzív

¹ Erre már az [1] alatt (164—175) rámutattunk.

számítások elvégzésére, mint a diadikus felbontáson alapuló egyéb módszerek-nél.²

c) Az alkalmazandó numerikus számítások mennyisége azonos nagyságrendű, mint a mátrix-számításon alapuló egyéb módszerek-nél.

A továbbiakban először az elemi transzformációval foglalkozunk, majd megmutatjuk, hogyan lehet felhasználni az elemi transzformációkat valamely adott A mátrix egy speciális bázisfelbontására. Ez a felbontás teszi lehetővé a lineáris egyenletrendszerek megoldásának egyszerű tárgyalását. Ehhez kapcsolódik a mátrixok inverziója, amely lényegében több olyan egyenletrendszer szimultán megoldását jelenti, amelyek azonos együtthatómátrixszal bírnak. A módszer hatékonyságát néhány numerikus példa bemutatásával kívánjuk illusztrálni.

1. §. Az elemi transzformáció és a bázisfelbontás

Legyenek a $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$ vektorok az m -dimenziós lineáris tér egyik bázisának vektorai. Ismeretes, hogy ez esetben a tér bármely \mathbf{x} vektora felírható az

$$(1.1) \quad \mathbf{x} = x_1 \mathbf{b}_1 + \dots + x_i \mathbf{b}_i + \dots + x_m \mathbf{b}_m$$

lineáris forma segítségével, ahol az $x_1, \dots, x_i, \dots, x_m$ skalárok a kérdéses bázisra vonatkozó koordináták. Amint már említettük, a szimplex módszer alapvető mozzanata az az eljárás, amellyel a bázis egyik vektorát kicseréljük a tér valamely alkalmasan megválasztott \mathbf{a} vektorával. Ezt az eljárást nevezzük *elemi transzformációnak*. Annak szükséges és elégséges feltétele, hogy a \mathbf{b}_i bázisvektort kicserélhessük a kérdéses \mathbf{a} vektorral, az, hogy az \mathbf{a} -nak a \mathbf{b}_i -re vonatkozó koordinátája 0-tól különböző legyen, vagyis az

$$(1.2) \quad \mathbf{a} = a_1 \mathbf{b}_1 + \dots + a_i \mathbf{b}_i + \dots + a_m \mathbf{b}_m$$

lineáris kombinációban teljesüljön az $a_i \neq 0$ követelmény.³ Ha (1.2)-ből kifejezzük \mathbf{b}_i -t, és az így nyert formulát behelyettesítjük a (1.1)-be, az

$$\mathbf{x} = (x_1 - \delta_i a_1) \mathbf{b}_1 + \dots + \delta_i \mathbf{a} + \dots + (x_m - \delta_i a_m) \mathbf{b}_m$$

kifejezést nyerjük, ahol

$$\delta_i = \frac{x_i}{a_i}.$$

E kifejezésből az új bázisra vonatkozó koordináták közvetlenül kiolvashatók. A szükséges számításokat célszerű az alábbi elrendezés szerint elvégezni⁴

² Mint pl. a [3], [4] és [5] alatt tárgyalt módszerek-nél.

³ Lásd az [1]-et a 101. oldalon!

⁴ A bázisból kilépő, ill. a bázisba belépő vektort egy-egy nyíllal jelöltük meg.

I.			II.		
Bázis	\mathbf{x}	\mathbf{a}	Bázis	\mathbf{x}	\mathbf{a}
\mathbf{b}_1	x_1	a_1	\mathbf{b}_1	$x_1 - \delta_i a_1$	0
\mathbf{b}_2	x_2	a_2	\mathbf{b}_2	$x_2 - \delta_i a_2$	0
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
$\leftarrow \mathbf{b}_i$	x_i	a_i	$\rightarrow \mathbf{a}$	δ_i	1
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
\cdot	\cdot	\cdot	\cdot	\cdot	\cdot
\mathbf{b}_m	x_m	a_m	\mathbf{b}_m	$x_m - \delta_i a_m$	0

A transzformációt meghatározó a_i skalárt *generáló elemnek* nevezzük. Az említett feltétel értelmében *generáló elem csak 0-tól különböző szám lehet.*

Tegyük fel ezután, hogy az

$$(1.4) \quad \mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]_{(m, n)}$$

mátrix oszlopvektorai a $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$ bázisban megadott vektorok. Ezzel kapcsolatban tűzzük ki az alábbi feladatot: Az adott bázis vektorait cseréljük ki, az \mathbf{A} oszlopvektoraival!

E feladat több egymás után végrehajtott elemi transzformációval oldható meg. Az egyes transzformációs lépéseknél a generáló elemet úgy kell megválasztani, hogy mindig valamely \mathbf{b}_i vektor cseréltessék ki egy lehetséges \mathbf{a}_j vektorral. Az eljárást mindaddig folytatjuk, amíg találunk az előbbi követelménynek eleget tevő — 0-tól különböző — generáló elemet. Az utolsó bázisban található \mathbf{a}_j vektorok szükségképpen bázisát alkotják az \mathbf{A} oszlopvektorai által meghatározott altérnek, az ún. *oszlopvektortérnek*. S mivel a mátrix rangja megegyezik az oszlopvektortér dimenziójával, a bázisba beépíthető \mathbf{a}_j vektorok maximális száma egyúttal az \mathbf{A} rangját is jelenti. Maga az egész eljárás nem tételezi fel eleve a rang ismeretét, ellenkezőleg: algoritmusként használható a rang meghatározására.

Tegyük fel, hogy az \mathbf{A} rangja r , és az előbb vázolt eljárás alapján éppen az első r

$$(1.5) \quad \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$$

oszlopvektor került be a bázisba. Ez nem sérti a tárgyalás általánosságát, hiszen — mint ismeretes — a rang független a vektorok sorrendjétől.⁵ Mivel az \mathbf{A} bármelyik oszlopvektora kifejezhető az oszlopvektortér bázisvektorainak lineáris kombinációjaként, azért

$$(1.6) \quad \mathbf{a}_j = \mathbf{A}_1 \mathbf{k}_j \quad (j = 1, 2, \dots, n),$$

⁵ Megjegyezzük azonban, hogy a számítások gyakorlati végrehajtásánál nem kell tekintettel lennünk a vektorok sorrendjére. Itt csak a tárgyalás egyszerűsítése végett folyamodtunk ehhez a feltevéshez.

ahol

$$(1.7) \quad \mathbf{A}_1 = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r],$$

a \mathbf{k}_j pedig olyan vektor, amelynek komponensei a j -edik oszlopvektornak az (1.5) alatti bázisra vonatkozó koordinátái. Az (1.6) és az (1.7) alapján

$$\mathbf{A} = [\mathbf{A}_1 \mathbf{k}_1, \mathbf{A}_1 \mathbf{k}_2, \dots, \mathbf{A}_1 \mathbf{k}_n] = \mathbf{A}_1 [\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_n] = \mathbf{A}_1 \mathbf{A}_2.$$

A jelen esetben

$$\mathbf{A}_2 = [\mathbf{E}_r, \mathbf{D}],$$

ahol az \mathbf{E}_r egy r -edrendű egységmatrix. Ezért

$$(1.8) \quad \mathbf{A} = \mathbf{A}_1 [\mathbf{E}_r, \mathbf{D}].$$

Ezzel az \mathbf{A} matrix egy speciális bázisfelbontásához jutottunk⁶. Az elnevezés azon a könnyen belátható tényen alapszik, hogy az \mathbf{A}_2 sorvektorai meg az \mathbf{A} sorvektorai által meghatározott altérnek, az ún. *sorvektortérnek*, adják meg egy bázisát.⁷ Egy adott numerikus példában a \mathbf{D} matrix elemei az utolsó elemi transzformáció végrehajtása után nyert táblázatból közvetlenül kiolvashatók.

Példaképpen tekintsük az

$$(1.9) \quad \mathbf{A} = \begin{bmatrix} 2 & 1 & 1 & 3 & 2 \\ 1 & -6 & 1 & 3 & 5 \\ 3 & 2 & 3 & 4 & 5 \\ -1 & 2 & -1 & -2 & -3 \end{bmatrix}$$

mátrixot. A számítások elvégzéséhez az alábbi táblázat szerint három elemi transzformáció szükséges.

Bázis	\mathbf{a}_1	\mathbf{a}_2	\mathbf{a}_3	\mathbf{a}_4	\mathbf{a}_5
$\leftarrow \mathbf{b}_1$	2	1	1	3	2
\mathbf{b}_2	1	-6	1	3	5
\mathbf{b}_3	3	2	3	4	5
\mathbf{b}_4	-1	2	-1	-2	-3
$\rightarrow \mathbf{a}_2$	2	1	1	3	2
\mathbf{b}_2	13	0	7	21	17
$\leftarrow \mathbf{b}_3$	-1	0	1	-2	1
\mathbf{b}_4	-5	0	-3	-8	-7

⁶ Megjegyezzük, hogy egy mátrixnak számtalan sok bázisfelbontása lehetséges. Egy adott $\mathbf{A}_1, \mathbf{A}_2$ bázisfelbontásból azonban az $(\mathbf{A}_1 \mathbf{M})(\mathbf{M}^{-1} \mathbf{A}_2)$ formula alapján bármely más bázisfelbontás előállítható. Az \mathbf{M} egy alkalmasan megválasztott nem szinguláris kvadratikus mátrixot jelent. (Lásd: [3] p. 431.)

⁷ Ugyanekkor az \mathbf{A}_1 sorvektorai az \mathbf{A} sorvektorainak az \mathbf{A}_2 által meghatározott bázisra vonatkozó koordinátáit adják.

$$(1.10) \quad \begin{array}{c|ccccc} \text{Bázis} & \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_5 \\ \hline \mathbf{a}_2 & 3 & 1 & 0 & 5 & 1 \\ \mathbf{b}_2 & 20 & 0 & 0 & 35 & 10 \\ \rightarrow \mathbf{a}_2 & -1 & 0 & 1 & -2 & 1 \\ \leftarrow \mathbf{b}_4 & -8 & 0 & 0 & -14 & \boxed{-4} \\ \hline \mathbf{a}_2 & 1 & 1 & 0 & 1,5 & 0 \\ \mathbf{b}_2 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{a}_3 & -3 & 0 & 1 & -5,5 & 0 \\ \rightarrow \mathbf{a}_5 & 2 & 0 & 0 & 3,5 & 1 \end{array}$$

További elemi transzformációra azért nem kerülhet sor, mert a \mathbf{b}_2 -nek megfelelő sorban már nem tudunk választani generáló elemet.

Az elmondottak alapján a

$$\left[\begin{array}{ccccc} 2 & 1 & 1 & 3 & 2 \\ 1 & -6 & 1 & 3 & 5 \\ 3 & 2 & 3 & 4 & 5 \\ -1 & 2 & -1 & -2 & -3 \end{array} \right] = \left[\begin{array}{ccc} 1 & 1 & 2 \\ -6 & 1 & 5 \\ 2 & 3 & 5 \\ 2 & -1 & -3 \end{array} \right] \left[\begin{array}{ccccc} 1 & 1 & 0 & 1,5 & 0 \\ -3 & 0 & 1 & -5,5 & 0 \\ 2 & 0 & 0 & 3,5 & 1 \end{array} \right]$$

bázisfelbontást nyerjük. Ha az oszlopvektorok sorrendjét történetesen az $\mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_5, \mathbf{a}_1, \mathbf{a}_4$ sorrendnek megfelelően választottuk volna meg, akkor a felbontást pontosan az (1.8) alatti formulának megfelelően kaptuk volna meg.

A számításokat nagymértékben leegyszerűsíthetjük, ha érvényesítjük az alábbi szempontokat.

- Az eredeti bázis vektorainak szimbólumát nem tüntetjük fel.
- A bázisba belépő vektorokhoz tartozó oszlopokat rendre elhagyjuk.
- Az egymás után következő táblázatokat nem egymás alá, hanem egymás mellé írjuk.

Ezen az alapon az (1.9) alatti táblázat a következővel helyettesíthető:

$$(1.11) \quad \begin{array}{c|ccccc|ccccc|ccccc|cc} & \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_5 & \mathbf{a}_1 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_5 & \mathbf{a}_1 & \mathbf{a}_4 & \mathbf{a}_5 & \mathbf{a}_1 & \mathbf{a}_4 \\ \hline \mathbf{a}_2 & 2 & \boxed{1} & 1 & 3 & 2 & 2 & 1 & 3 & 2 & 3 & 5 & 1 & 1 & 1,5 \\ & 1 & -6 & 1 & 3 & 5 & 13 & 7 & 21 & 17 & 20 & 35 & 10 & 0 & 0 \\ \mathbf{a}_3 & 3 & 2 & 3 & 4 & 5 & -1 & \boxed{1} & -2 & 1 & -1 & -2 & 1 & -3 & -5,5 \\ \mathbf{a}_5 & -1 & 2 & -1 & -2 & -3 & -5 & -3 & -8 & -7 & -8 & -14 & \boxed{-4} & 2 & 3,5 \end{array}$$

2. §. Lineáris egyenletrendszerek megoldása

Tekintsük az

$$(2.1) \quad \mathbf{A} \mathbf{x} = \mathbf{c}$$

egyenletrendszert. Ezt az egyenletrendszert megoldani annyit jelent, mint meghatározni mindazon $\mathbf{x}^* = [x_1, x_2, \dots, x_n]$ pontok halmazát, amelyek kielégítik a (2.1) alatti követelményt. Ha e halmaz nem üres, az egyenletrendszert *kompatibilisnek* nevezzük, ellenkező esetben *inkompatibilisnek*. Kézen fekvő,

hogy a kompatibilitás ekvivalens azzal a feltétellel, hogy a \mathbf{c} benne fekszik az \mathbf{A} oszlopvektortérben.

Tegyük fel, hogy egyenletrendszerünk kompatibilis, továbbá: az \mathbf{A} az (1.8)-nak megfelelően az

$$(2.2) \quad \mathbf{A} = \mathbf{A}_1[\mathbf{E}_r, \mathbf{D}]$$

alakban bontható fel. A kompatibilitás követelménye értelmében a \mathbf{c} is kifejezhető az \mathbf{A}_1 által meghatározott bázisban, azaz

$$(2.3) \quad \mathbf{c} = \mathbf{A}_1 \mathbf{d}.$$

A (2.2) és a (2.3) alapján azonban a (2.1) így írható fel:

$$(2.4) \quad \mathbf{A}_1[\mathbf{E}_r, \mathbf{D}] \mathbf{x} = \mathbf{A}_1 \mathbf{d}.$$

Mivel az \mathbf{A}_1 oszlopvektorai lineárisan függetlenek a (2.1), illetve a (2.4) alatti egyenletrendszer ekvivalens az

$$(2.5) \quad [\mathbf{E}_r, \mathbf{D}] \mathbf{x} = \mathbf{d}$$

egyenletrendszerrel, amely már könnyen megoldható. Bontsuk fel ugyanis az \mathbf{x} vektort az \mathbf{x}_1 és \mathbf{x}_2 szimbolumokkal jelölt blokkokra, ahol az \mathbf{x}_1 komponenseinek száma r , az \mathbf{x}_2 komponenseinek száma pedig $n - r = s$. Ez esetben a (2.5) helyett az

$$[\mathbf{E}_r, \mathbf{D}] \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \mathbf{d}$$

egyenlőséget nyerjük, ahonnan

$$(2.6) \quad \mathbf{x}_1 = \mathbf{d} - \mathbf{D} \mathbf{x}_2.$$

Ez pedig éppen az egyenletrendszer megoldását jelenti a szokásos formában. Az \mathbf{x}_1 komponensei az ún. *kötött ismeretlenek*, az \mathbf{x}_2 komponensei meg az ún. *szabad ismeretlenek*. Mivel az \mathbf{x}_2 komponensei egymástól függetlenül bármilyen valós értéket felvehetnek, az s skalárt az *egyenletrendszert szabadságfokának* nevezzük.

A (2.6) alatti megoldást olyan formában is felírhatjuk, amely — úgy véljük — mélyebb bepillantást enged a dolog lényegébe. Csatoljuk ugyanis a (2.6)-hoz a triviálisan teljesülő

$$\mathbf{x}_2 = \mathbf{0} + \mathbf{E}_s \mathbf{x}_2$$

egyenlőséget, ahol az \mathbf{E}_s egy s -edrendű egységmátrix. Ezzel az

$$\mathbf{x} = \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{D} \\ \mathbf{E}_s \end{bmatrix} \mathbf{x}_2$$

kifejezéshez jutunk. Ha az \mathbf{x}_2 helyébe, amelynek komponensei úgyis tetszőlegesen, a \mathbf{t} vektort írjuk, a megoldást az

$$(2.7) \quad \mathbf{x} = \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{D} \\ \mathbf{E}_s \end{bmatrix} \mathbf{t} = \mathbf{x}_0 + \mathbf{B} \mathbf{t}$$

formában kapjuk meg. E formulában a \mathbf{t} helyébe bármilyen s komponensű vektort behelyettesíthetünk. Ebből viszont az következik, hogy az \mathbf{x}_0 a $\mathbf{t} = \mathbf{0}$ vektorhoz tartozó *partikuláris megoldást* jelenti, a $\mathbf{B} \mathbf{t}$ pedig a \mathbf{B} oszlopvektor-

terének valamelyik vektorát. Innen könnyen adódik az ismert tétel, amely szerint a $\mathbf{B} \mathbf{t}$ alakú vektorok egy, az \mathbf{A} sorvektorterére merőleges alteret alkotnak, amelynek a dimenziója s . Az állítás első része az

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{A}_1 [\mathbf{E}_r, \mathbf{D}] \begin{bmatrix} -\mathbf{D} \\ \mathbf{E}_s \end{bmatrix} = \mathbf{0}$$

összefüggésből következik, a második része pedig abból a tényből, hogy a \mathbf{B} olyan trapéz alakú mátrix, amelynek egyik blokkja egy s -edrendű egység-mátrix.

Összefoglalva: a (2.3) alatti egyenletrendszer összes lehetséges megoldásai egy olyan s dimenziós hipersíkot állítanak elő, amely merőleges az együttható-mátrix sorvektorterére.

A kérdéses most már az, hogy egy adott numerikus példában hogyan kaphatjuk meg a megoldásban szereplő állandókat. A \mathbf{D} mátrixot illetőleg már láttuk, hogy az — elemi transzformációk segítségével — r lépésben meghatározható. Nyilvánvaló azonban, hogy ugyanezek a transzformációk a \mathbf{d} vektor meghatározására is alkalmasak. Semmi mást nem kell tenni, mint az \mathbf{A} mátrixot kibővíteni a \mathbf{c} oszlopvektorral és a szükséges számításokat ezzel a *bővített mátrix*sal elvégezni. A \mathbf{c} koordinátái közül természetesen nem választ-hatunk generáló elemet. Az r -edik transzformációs lépés után kiderül, hogy a \mathbf{c} benne fekszik-e az \mathbf{A} oszlopvektorterében vagy nem, azaz: kompatibilis-e az egyenletrendszer vagy sem. A számítások tehát a rang, illetve a bázis kérdése mellett a kompatibilitás kérdését is automatikusan felderítik.

A kompatibilitást illetőleg egyébként triviálisak az alábbi állítások.

a) A homogén egyenletrendszer (amikor tehát $\mathbf{c} = \mathbf{0}$) mindig kompatibilis.

b) Ha $r = m$ az egyenletrendszer ugyancsak kompatibilis.

A szabadságfok tekintetében gyakorlati szempontból külön figyelmet érdemel az az eset, amikor $s = 0$. Ekkor, amennyiben az egyenletrendszer kompatibilis, a megoldás egyetlen pontra redukálódik. (Az egyenletrendszer egyértelműen oldható meg.)

A numerikus eljárás illusztrálására megoldjuk a következő egyenlet-rendszert:

$$\mathbf{A} \mathbf{x} = \begin{bmatrix} 7 \\ -4 \\ 18 \\ -2 \end{bmatrix},$$

ahol \mathbf{A} az (1.9)-ben szereplő mátrix. A számítások végrehajtása tehát az ottani-tól csupán abban fog különbözni, hogy az \mathbf{A} oszlopvektorai mellett a táblázatban a \mathbf{c} vektort is feltüntetjük. Az új táblázatban — a numerikus igények-nek megfelelően — a vektorszimbólumok helyett mindenütt a megfelelő ismeretlenek szimbólumát fogjuk feltüntetni.

	x_1	x_2	x_3	x_4	x_5		x_1	x_3	x_4	x_5		x_1	x_4	x_5		x_1	x_4	
x_2	2	<u>1</u>	1	3	2	7	1	1	3	2	7	3	5	1	3	1	1,5	2
	1	-6	1	3	5	-4	13	7	21	17	38	20	35	10	10	0	0	0
x_3	3	2	3	4	5	18	-1	<u>1</u>	-2	1	4	-1	-2	1	4	-3	-5,5	3
x_5	-1	2	-1	-2	-3	-2	-5	-3	-8	-7	-16	-8	-14	<u>-4</u>	-4	2	3,5	1

A (2.6)-nak megfelelően a megoldás tehát a következő:

$$\begin{aligned}x_2 &= 2 - 1x_1 - 1,5x_4 \\x_3 &= 3 + 3x_1 + 5,5x_4 \\x_5 &= 1 - 2x_1 - 3,5x_4\end{aligned}$$

Ez a (2.7) szerinti formában:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 3 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & -1,5 \\ 3 & 5,5 \\ 0 & 1 \\ -2 & -3,5 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$$

A bemutatott eljárás jól alkalmazható a *mátrixok inverzének* meghatározására is. Egy adott **A** mátrix jobboldali inverzén például, mint ismeretes, azt az **X** mátrixot értjük, amely eleget tesz az

$$\mathbf{A}\mathbf{X} = \mathbf{E}$$

követelményeknek. Az **X** meghatározása az azonos **A** együtthatómátrixszal rendelkező

$$\mathbf{A}\mathbf{x}_i = \mathbf{e}_i \quad (i = 1, 2, \dots, m)$$

alakú egyenletrendszerek szimultán megoldására vezethető vissza, ahol az \mathbf{x}_i az **X** mátrix, az \mathbf{e}_i pedig az **E** mátrix i -edik oszlopvektora. Ebből következik, hogy eljárásunkat csupán annyiban kell módosítanunk, hogy a **c** vektor helyébe az **E** oszlopvektorait kell írunk. Könnyen belátható, hogy a jobboldali inverz akkor és csak akkor létezik, ha az **A** sorvektorai lineárisan függetlenek.

Határozzuk meg példaképpen az

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 3 \\ 2 & 1 & 0 & 1 \\ 3 & 1 & 1 & 1 \end{bmatrix}$$

jobboldali inverzét. A megoldást az alábbi táblázatból olvashatjuk ki:

	\mathbf{a}_1	\mathbf{a}_2	\mathbf{a}_3	\mathbf{a}_4		\mathbf{a}_2	\mathbf{a}_3	\mathbf{a}_4		\mathbf{a}_3	\mathbf{a}_4		\mathbf{a}_4	
\mathbf{a}_1	<u>1</u>	1	1	3	1 0 0	1	1	3	1 0 0	-1 -2	-1	1 0	-1	-0,5 0 0,5
\mathbf{a}_2	2	1	0	1	0 1 0	<u>-1</u>	-2	-5	-2 1 0	2 5	2	-1 0	3	1 1 -1
\mathbf{a}_3	3	1	1	1	0 0 1	-2	-2	-8	-3 0 1	<u>2</u>	2	1 -2 1	1	0,5 -1 0,5

E szerint a keresett inverz:

$$\begin{bmatrix} -0,5 & 0 & 0,5 \\ 1 & 1 & -1 \\ 0,5 & -1 & 0,5 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -3 \\ -1 \\ 1 \end{bmatrix} t$$

ahol a t bármilyen skalár lehet.

Numerikus szempontból érdeklődésre tarthat talán számot, hogy a számításokban további egyszerűsítéseket is eszközölhetünk. Ez abban áll, hogy az induló táblázatban nem írjuk fel az **E** oszlopvektorait. Az inverz mátrixban szereplő vektorok csak lépésről lépésre kerülnek be a táblázatba — az eredeti mátrixnak a bázisba bekerülő vektorai helyébe. Ha tehát az inverz vektorait rendre az i_1, i_2, \dots szimbólumokkal jelöljük, akkor az előbbi táblázat helyett a következőt alkalmazhatjuk:

	a_1	a_2	a_3	a_4	i_1	a_2	a_3	a_4	i_1	i_2	a_3	a_4	i_1	i_2	i_3	a_4
a_1	1	1	1	3	1	1	1	3	-1	1	-1	-2	-0,5	0	0,5	-1
a_2	2	1	0	1	-2	-1	-2	-5	2	-1	2	5	1	1	1	3
a_3	3	1	1	1	-3	-2	-2	-8	1	-2	1	2	0,5	-1	0,5	1

(Beérkezett: 1958. február 15.)

IRODALOM

- [1] KREKÓ — BACSKAY: *Bevezetés a lineáris programozásba*. Közgazdasági és Jogi Könyvkiadó. Budapest, 1957.
- [2] DANTZIG, G. B.: "Maximation of a linear function of variables subject to linear inequalities." *Koopmans: Activity analysis of production and allocation*. John Wiley. New York, 1951. p. 339—347.
- [3] EGERVÁRY J.: „Mátrixfüggvények kanonikus előállításáról és annak néhány alkalmazásáról”. *A Magyar Tudományos Akadémia III. Osztályának Közleményei* 3 (1953) 417—458.
- [4] EGERVÁRY J.: „Mátrixok diadikus előállításán alapuló módszer bilineáris alakok transzformációjára és lineáris egyenletrendszerek megoldására” *A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei* 2 (1953) 11—32.
- [5] EGERVÁRY J.: „Régi és új módszerek lineáris egyenletrendszerek megoldására.” *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* 1 (1956) 109—123.

РЕШЕНИЕ СИСТЕМ ЛИНЕЙНЫХ УРАВНЕНИЙ МЕТОДОМ СИМПЛЕКСА

В. КРЕКО

Резюме

Автор показывает, что несколько измененный метод симплекса G. B. DANTZIG-а может применяться и для решения систем линейных уравнений. Основа метода — так называемое элементарное преобразование. Это такое специальное линейное преобразование, с помощью которого из данного базиса можно перейти в такой новой базис, который отличается от исходного лишь в одном векторе.

Рассмотрим некоторую систему уравнений, заданную в виде

$$Ax = b.$$

Предположим, что ранг \mathbf{A} равен r и что векторы, стоящие в её первых r столбцах, линейно независимы. В этом случае \mathbf{A} может быть факторизованно по образцу соотношения

$$\mathbf{A} = \mathbf{A}_1[\mathbf{E}_r \mathbf{D}],$$

где \mathbf{A}_1 матрица, образованная из первых r столбцов \mathbf{A} , а \mathbf{E}_r некоторая единичная матрица порядка r . Если система уравнений разрешима, то \mathbf{b} может быть записан и в виде

$$\mathbf{b} = \mathbf{A}_1 \mathbf{d}.$$

Но тогда исходная система уравнений эквивалентна системе уравнений

$$[\mathbf{E}_r, \mathbf{D}] \mathbf{x} = \mathbf{d},$$

общее решение которой

$$\mathbf{x} = \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{D} \\ \mathbf{E}_s \end{bmatrix} \mathbf{t}.$$

Здесь \mathbf{E}_s единичная матрица порядка s а \mathbf{t} любой вектор порядка s . s — степень свободы системы уравнений.

С помощью r элементарных преобразований как \mathbf{D} , так и \mathbf{d} могут быть численно определены. Метод не предполагает заранее ни знания ранга, ни разрешимости.

С помощью этого метода может быть просто произведено и вычисление обратных матриц.

SOLUTION OF A SYSTEM OF LINEAR EQUATIONS WITH THE SIMPLEX METHOD

by

B. KREKÓ

Abstract

This paper shows that the simplex method due to G. B. DANTZIG is convenient — with some modification — to solve a set of simultaneous linear equations. The essence of the procedure is the so-called elementary transformation. It is a special linear transformation by the means of which we can go from a given basis to another basis differing only in a single vector.

Let us consider a system of linear equations in the following form:

$$\mathbf{A} \mathbf{x} = \mathbf{b}.$$

We suppose that the rank of \mathbf{A} is r and — accidentally — the first r column — vectors of \mathbf{A} are linear independent. In this case

$$\mathbf{A} = \mathbf{A}_1 \cdot [\mathbf{E}_r, \mathbf{D}],$$

where \mathbf{A}_1 is a matrix containing the first r column-vector of \mathbf{A} and \mathbf{E}_r is a unit matrix of order r . If the system of equations is compatible, we can write

$$\mathbf{b} = \mathbf{A}_1 \mathbf{d}.$$

In these circumstances the original system is equivalent with the system

$$[\mathbf{E}_r, \mathbf{D}] \mathbf{x} = \mathbf{d},$$

having the general solution

$$\mathbf{x} = \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{D} \\ \mathbf{E}_s \end{bmatrix} \mathbf{t},$$

where \mathbf{E}_s is a unit matrix of order s and \mathbf{t} is an arbitrary vector with s components. s is the degree of freedom of the system.

With the help of r elementary transformations we can numerically determine both \mathbf{D} and \mathbf{d} .

The procedure does not suppose the knowledge of the rank and the compatibility.

The procedure presented is also convenient to determine the inverse of a matrix nonsingular.



ELLENÁRAMÚ SZÉTVALASZTÓ VEGYIPARI ALAPMŰVELETEK ELMÉLETI FOKOZATSZÁMÁNAK MEGHATÁROZÁSÁRÓL, II.

RÓZSA PÁL és SÁRKÁNY György¹

E dolgozatban — ahhoz hasonlóan, ahogyan azt az előző közleményünkben [1] tettük —, a középén táplált reflux nélküli extraktornak (mint általános alpműveleti modellnek), az elméleti fokozatszámára kívánunk — bizonyos közelítő feltevések alkalmazásával — számításra alkalmas explicit képleteket nyerni. Az előző dolgozatunkban feltételeztük, hogy az egymással szemben áramló két fázis mennyiségi sebessége állandó és kiszámítható sebességjellegű paraméterrel jellemezhető. Jelen dolgozatunkban ettől a feltételtől eltekintünk, tehát azt az *általánosabb* esetet vizsgáljuk, amikor az egyes fokozatokon szereplő koncentrációk mellett a *fázisok változó mennyiségi sebességei* (L_1 és V_1) is ismeretlenek.

Az első fejezetben megadjuk azokat a részben már előző dolgozatunkban is felhasználott, részben pedig ott nem szereplő alapvető összefüggéseket (anyagmérlegeket, illetve hőmérleget, valamint az egyes fázisok koncentrációi közötti egyensúlyi összefüggéseket), amelyek alapján a következő fejezetben a fokozatszám meghatározására szükséges egyenletrendszereket írjuk fel. Az egyenletrendszerek felírásánál most is azt az elvet követjük, hogy a tapasztalati összefüggésekből nyert görbéket húrpoligonokkal helyettesítjük. A fokozatszám meghatározása úgy történik, hogy az egyenletrendszereket a húrpoligonoknak megfelelően választott csoportokra osztjuk, s egy-egy egyenletcsoport egyenleteinek számát határozzuk meg. A feladat matematikailag nemlineáris egyenletrendszerre vezet, az ismeretlenekre pedig tört-lineáris rekurziós formulát nyerünk. Ennek megoldásával a tulajdonképpeni feladat — tehát az egyenletek *számának* a meghatározása — is megoldhatóvá válik. Az eredményül kapott számításra alkalmas képletek a 292. oldalon találhatók (27), (28), (28'). (E képletekben szereplő közvetett jelölések jelentését a dolgozat végén feltüntetett jegyzék alapján találhatjuk meg.) Végül (középső táplálás nélküli) extraktor, refluxos extraktor és rektifikáló esetére megadjuk azokat a specifikáló feltételeket, amelyek segítségével az általános eredmény ezekre alkalmazható.

1. §. Alapvető összefüggések

Előző tanulmányunkban [1] a rektifikálást és abszorpciót egyaránt tágabb értelmű extrakciónak fogtuk fel és ezen alapon úgynevezett általános extraktorra vezettük vissza azokat, valamint az extrakció különböző változa-

¹ Budapesti Műszaki Egyetem Vegyipari Műveletek és Gépek tanszék.

tait. Az általános extraktor (és a belőle specifikált speciális extraktor, rektifikáló és abszorber) elméleti fokozatszámának kiszámítására explicit közelítő képleteket vezettünk le.

A számítás az egyes elméleti fokozatokra felírt kétféle egyenlet felhasználásával történt:

a) az értékes komponens anyagmérlege,

b) az értékes anyag megoszlási (fázisegyensúlyi) egyenlete. Az így nyert n fokozatra vonatkozó $2n$ egyenletből az egyenletek számát kellett meghatározni.

Számításainkban feltételeztük, hogy az egymással szemben áramló két fázis mennyiségi sebessége (alkalmasan választott dimenzió esetén) állandó moláris sebességjellegű paraméterrel jellemezhető, aminek következtében a C értékes komponens anyagmérlege

$$(a) \quad -L_{l-1}x_{l-1} + L_l x_l + V_l y_l - V_{l+1}x_{l+1} = T x_T \delta_{lk}, \quad (l = 1, 2, \dots, n)$$

az $L = \text{konst}$ és $V = \text{konst}$ feltételeknek megfelelően egyszerűsödik.

Az így nyert n egyenletből álló egyenletrendszer kiegészítette a C anyag fázisegyensúlyi görbáját helyettesítő húrpoligon, amely az x_l és y_l mennyiségek között ad összefüggést. (Lásd a 282. oldalon az (f) egyenletrendszert.)

A gyakorlatban L és V extrakciónál rendszerint, rektifikálásnál pedig sok esetben *nem állandó*. (Lásd a 280. oldal ² lábjegyzetét.) Ennek az a következménye, hogy az egyes fokozatokon szereplő x_l és y_l koncentrációk mellett a fázisok mennyiségi sebességei (L_l és V_l) is ismeretlenek. Ezért az egyenletek számának meghatározásához további, a viszonylagos fázismennyiséget meghatározó egyenletrendszerekre is szükség van.

Ha L és V állandóságának feltételezése nem jogosult, akkor extrakciónál a B anyag mérlegét, rektifikálásnál a fázismennyiséget meghatározó entalpiák (H és h) mérlegét, valamint a fokozatonként változó L_l és V_l (kmól fázis/óra) mérlegeit is fel kell írni minden fokozatra. Így újabb $2n$ egyenletet nyerünk, és ezt az újabb egyenletrendszert az előbbi kettőhöz kell csatolni.

Extrakciónál a B oldószernek a fázisegyensúlyban való részvételét az úgynevezett oldószermentes alapon (empirikusan) megszerkesztett diagram. rektifikálásnál pedig az ugyancsak empirikus entalpia-koncentráció diagram tünteti fel. (Tárgyalásaink abszorpcióra nem vonatkoznak, mert ott a fázismennyiség változását más természetű jelenség szabja meg.)

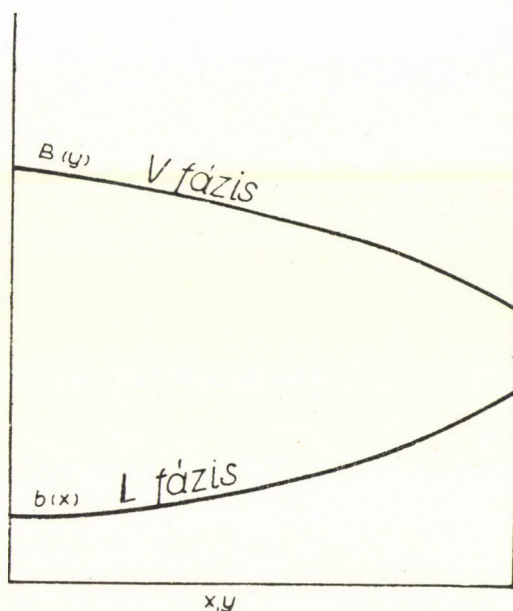
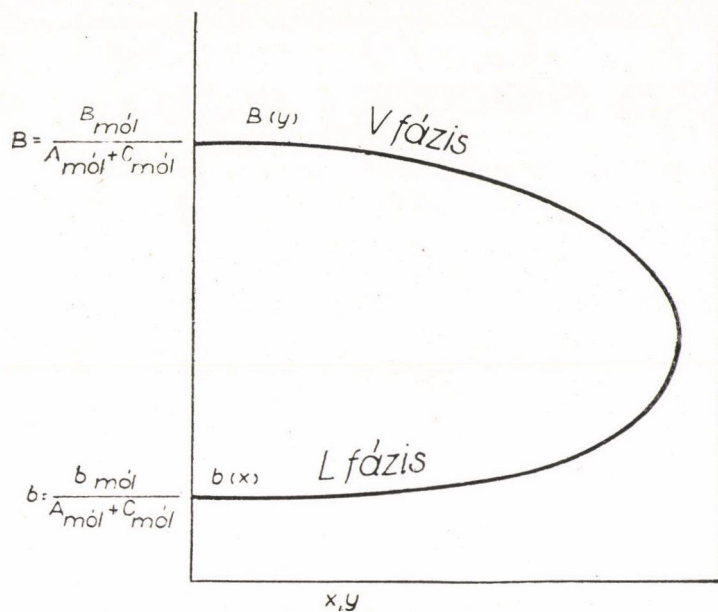
Ha abszcisszául ismét az

$$x = \left(\frac{C_{mól}}{A_{mól} + C_{mól}} \right) L \text{ fázisban} \quad \text{és} \quad y = \left(\frac{C_{mól}}{A_{mól} + C_{mól}} \right) V \text{ fázisban}$$

definíciókat fogadjuk el, akkor a B diagram két tipikus alakját az 1. ábra mutatja.

Ennek megfelelően a hiányzó, a fázismennyiséget meghatározó két egyenletrendszer extrakciónál a következő:

$$(b) \quad \begin{aligned} -L_{l-1}b_{l-1} + L_l b_l + V_l B_l - V_{l+1}B_{l+1} &= T b_T \delta_{lk}, & (l = 1, 2, \dots, n) \\ -L_{l-1} + L_l + V_l - V_{l+1} &= T \delta_{lk}, & (l = 1, 2, \dots, n) \end{aligned}$$



1. ábra

Ha az utóbbi egyenleteket l -től n -ig összeadjuk, azaz az extraktornak az l -edik fokozattól az n -edik fokozatig terjedő oszloprészére írjuk fel a teljes anyagmérleget, akkor a következő összefüggéseket nyerjük:

$$(c) \quad V_l = L_{l-1} + G_l, \quad (l = 1, 2, \dots, n),$$

ahol $G_l = T + V_0 - L_n$, ha $l \leq k$. Ha viszont $l > k$, akkor $G_l = V_0 - L_n$. Egységes tárgyalás érdekében bevezetjük az úgynevezett egységugrás függvényét, amelyet az alábbi összefüggéssel definiálunk:

$$\mathcal{E}_{lk} = \begin{cases} 1, & \text{ha } k - l \geq 0 \\ 0, & \text{ha } k - l < 0 \end{cases} \quad l = 1, 2, \dots, n.$$

Ezzel tehát

$$(d) \quad G_l = \mathcal{E}_{lk} T + V_0 - L_n,$$

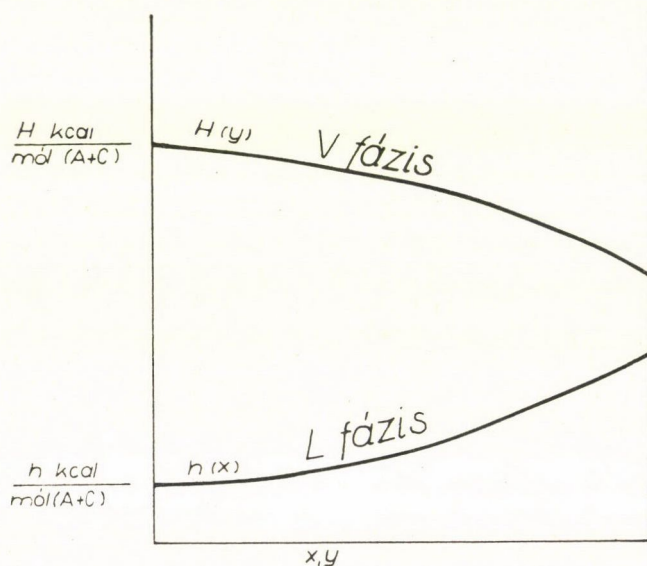
(k ugyanis rögzített érték).

Rektifikálásnál a (másik fázisképző) B oldószer szerepét a párolgási hő veszi át, amit az A és C anyag entalpia-koncentráció diagramja tüntet fel. (Merkel-diagram, 2. ábra). Ezen a diagramon $(H - h)_{x=0}$ az A anyag, $(H - h)_{x=1}$ pedig a C párolgáshőjét jelenti.

A diagram alapján írhatjuk fel rektifikálásnál a hiányzó két egyenletet:²

$$-L_{l-1}h_{l-1} + L_l h_l + V_l H_l - V_{l+1} H_{l+1} = Th_T \delta_{lk}, \quad (l = 1, 2, \dots, n)$$

$$-L_{l-1} + L_l + V_l - V_{l+1} = T \delta_{lk}, \quad (l = 1, 2, \dots, n).$$

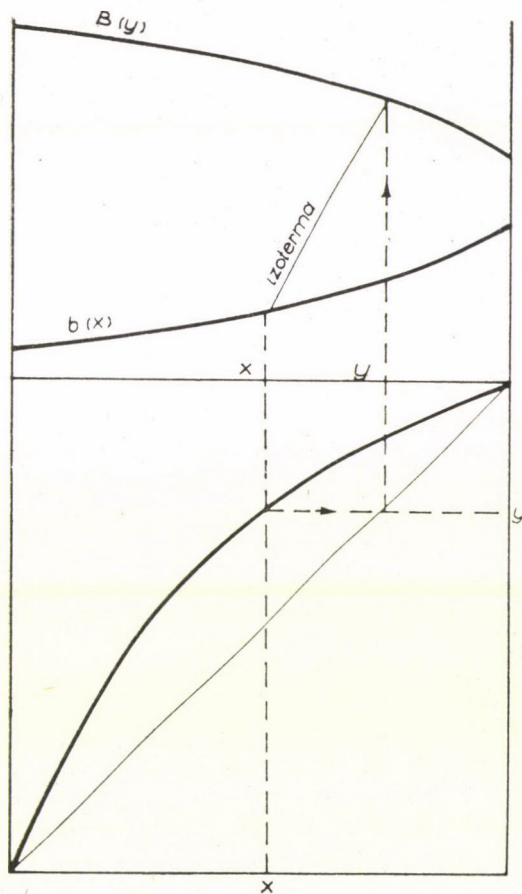


2. ábra

² L és V koncentráció-függetlensége akkor jelentkezik, ha az 1. ábra $B(y)$ és $b(x)$, illetve a 2. ábra $H(y)$ és $h(x)$ vonalai párhuzamos egyenesek.

Megjegyezzük, hogy $B(y)$ és $b(x)$, illetve $H(y)$ és $h(x)$ között a C anyag megoszlását meghatározó $y(x)$ egyensúlyi görbe (lásd az (f) egyenletrendszert a 282. oldalon) adja meg a kapcsolatot, az úgynevezett izotermák révén (3. ábra).

Most összesen négy egyenletrendszerünk van az egyenletek számának meghatározására (a, b, c, f). Az egyenletrendszerek megoldásához szükségesek



3. ábra

a $B(y)$, $b(x)$, illetve $H(y)$ és $h(x)$ függvények. Minthogy ezek csak empirikus grafikonok formájában ismeretesek, eddig csupán a megfelelő diagramok segítségével fokozatról fokozatra végzett próbálgatással, vagy pedig grafikus módszerrel lehetett a fokozatszámot meghatározni (SOREL módszere). (Lásd [2], 118. oldal.) A fokozatszámnak zárt képlettel való meghatározása — elméleti úton nyert $B(y)$, $b(x)$, illetve $H(y)$, $h(x)$ függvények híján — úgy is lehetséges, hogy a rendelkezésre álló empirikus grafikonokat megfelelő függvényekkel közelítjük, amivel újabb két egyenletrendszert nyerünk. A számítás most

is a k -adik fokozaton táplált általános extraktorra vonatkozik (lásd előző dolgozatunk 1. ábráját). A táplálék fizikai állapotát most annak b_T adata határozza meg (extrakciónál ez a mennyiség a táplálék B anyagtartalmát, rektifikálásnál h_T entalpiáját jelenti). Ez szabja meg T kettéoszlását L és V fázisok között. Az előző tanulmányunkban szereplő q mennyiség egyenrangú b_T -vel, illetve h_T -vel. Köztük q -nak ott nem említett definíciója adja meg az összefüggést:

$$(e) \quad q = \frac{B(x_T) - b_T}{B(x_T) - b(x_T)}.$$

Megjegyezzük még, hogy a tárgyalás egyszerűsítése érdekében feltételezzük, hogy L_0 és V_0 az oldószér-, illetve entalpiadiagram megfelelő görbéje által kijelölt állapotban vannak, azaz L_0 oldószertartalma

$$b_0 = b'_0 x_0 + b''_0,$$

illetve

$$h_0 = h'_0 x_0 + h''_0,$$

stb.

2. §. Az általános alpművelet egyenletrendszere

Amint az első fejezetben említettük és ahhoz hasonlóan, ahogy előző dolgozatunkban tettük, az elméleti fokozatszám meghatározására a fázis-egyensúlyi görbét és a viszonylagos fázismennyiséget meghatározó görbéket az (x_0, x_n) intervallumban alkalmasan választott x_{0i} ($i = 1, 2, \dots, v$) abszcisszájú pontokon áthaladó $v + 1$ lineáris szakaszból álló húrpolygonokkal közelítjük.³

Ennek értelmében $x_{00} \equiv x_0$ és $x_{0,v+1} \equiv x_n$. A fázisegyensúlyi görbe így módon nyert lineáris szakaszainak egyenlete legyen

$$(f) \quad y = K_i x + K'_i \quad (i = 0, 1, 2, \dots, v),$$

$$x_{0,i-1} \leq x \leq x_{0i}$$

az extrakció fázisviszonyait meghatározó görbék lineáris szakaszainak egyenletei:

$$(g) \quad B = B'_i y + B''_i \quad (i = 0, 1, 2, \dots, v)$$

és

$$(h) \quad b = b'_i x + b''_i \quad (i = 0, 1, 2, \dots, v)$$

$$x_{0,i-1} \leq x \leq x_{0i}$$

³ Rektifikálásnál az entalpia-vonalak egyenessel való közelítésének gondolata E. KIRSCHBAUMTÓL származik ([3] 154. o.). Ilyen alapon végzett elméleti fokozatszám meghatározást lásd [4]-ben.

(oldószermentes alapon felírt oldószertartalmak), rektifikálásnál pedig

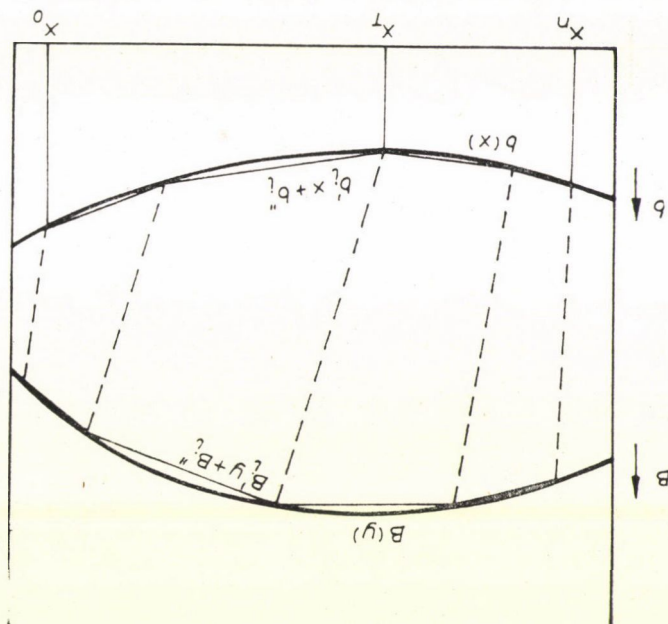
$$H = H'_i y + H''_i \quad (i = 0, 1, 2, \dots, v)$$

és

$$h = h'_i x + h''_i \quad (i = 0, 1, 2, \dots, v)$$

$$x_{0,i-1} \leq x \leq x_{0i}$$

(entalpia görbék). Itt az i -edik egyenletben y értelmezési tartományát az (f)-ből adódó határok szabják meg (4. ábra). Feladatunk abban áll, hogy az



4. ábra

egyes szakaszokhoz tartozó p_i ($i = 0, 1, 2, \dots, v$) elméleti fokozatszámot meghatározzuk. Az egyes fokozatokon szereplő koncentrációkat és anyagáramokat kettős indexszel fogjuk jelölni. A k -adik fokozatnak, ahol a táplálás történik, $(p_{k-1}, z-1)$ indexpár, az utolsó (n -edik) fokozatnak pedig a (p_v, v) indexpár fog megfelelni, és általában minden egyes lineáris szakaszhoz tartozó fokozatcsoport utolsó, azaz (p_i, i) indexpárral jellemzett fokozatához tartozó megfelelő koncentráció értékek megegyeznek a következő csoport 0 indexű — már előre megválasztott — koncentrációértékével:

$$x_{p_i, i} = x_{0, i+1}; \quad y_{p_i, i} = y_{0, i+1}, \quad (i = 0, 1, 2, \dots, v).$$

A megfelelő anyagáramok pedig (amelyek értékeit azonban előre nem ismerjük):

$$L_0 = L_{00}; \quad L_{p_i, i} = L_{0, i+1}; \quad V_{p_i, i} = V_{0, i+1}, \quad (i = 0, 1, 2, \dots, v).$$

Ennek alapján felírjuk a bevezetőben említett (a) egyenletrendszer. Az egyes fokozatcsoportokhoz tartozó első egyenletek:

$$(1) \left\{ \begin{array}{l} L_{10} x_{10} + V_{10} y_{10} - V_{20} y_{20} = L_0 x_0 \\ - L_{p_{i-1}, i-1} x_{p_{i-1}, i-1} + L_{1i} x_{1i} + V_{1i} y_{1i} - V_{2i} y_{2i} = 0; \quad i = 1, 2, \dots, v, \\ \text{az utolsó egyenletek:} \\ - L_{p_{i-1}, i} x_{p_{i-1}, i} + L_{pi} x_{pi} + V_{pi} y_{pi} - V_{1,i+1} y_{1,i+1} = \\ = \begin{cases} 0, & \text{ha } i = 0, 1, \dots, \kappa - 2, \kappa, \dots, v-1 \\ T x_T & \text{ha } i = \kappa - 1, \end{cases} \\ - L_{p_{v-1}, v} x_{p_{v-1}, v} + V_{p,v} y_{p,v} = V_0 y_0 - L_n x_n, \\ \text{a közbülső egyenletek pedig:} \\ - L_{j-1, i} x_{j-1, i} + L_{ji} x_{ji} + V_{ji} y_{ji} - V_{j+1, i} y_{j+1, i} = 0, \quad \begin{matrix} j = 2, 3, \dots, p_i - 1 \\ i = 0, 1, \dots, v. \end{matrix} \end{array} \right.$$

A (b) egyenletrendszerben az első egyenletek:

$$(2) \left\{ \begin{array}{l} L_{10}(b'_0 x_{10} + b''_0) + V_{10}(B'_0 y_{10} + B''_0) - V_{20}(B'_0 y_{20} + B''_0) = L_0(b'_0 x_0 + b''_0) \\ - L_{p_{i-1}, i-1}(b'_{i-1} x_{p_{i-1}, i-1} + b''_{i-1}) + L_{1i}(b'_i x_{1i} + b''_i) + V_{1i}(B'_i y_{1i} + B''_i) - \\ - V_{2i}(B'_i y_{2i} + B''_i) = 0 \quad i = 1, 2, \dots, v, \\ \text{az utolsó egyenletek:} \\ - L_{p_{i-1}, i}(b'_i x_{p_{i-1}, i} + b''_i) + L_{pi}(b'_i x_{pi} + b''_i) + V_{pi}(B'_i y_{pi} + B''_i) - \\ - V_{1,i+1}(B'_{i+1} y_{1,i+1} + B''_{i+1}) = \begin{cases} 0, & \text{ha } i = 0, 1, \dots, \kappa - 2, \kappa, \dots, v-1 \\ T(b'_{\kappa-1} x_T + b''_{\kappa-1}), & \text{ha } i = \kappa - 1 \end{cases} \\ - L_{p_{v-1}, v}(b'_v x_{p_{v-1}, v} + b''_v) + V_{p,v}(B'_v y_{p,v} + B''_v) = Q' - L_n(b'_v x_n + b''_v), \\ \text{a közbülső egyenletek pedig:} \\ - L_{j-1, i}(b'_i x_{j-1, i} + b''_i) + L_{ji}(b'_i x_{ji} + b''_i) + V_{ji}(B'_i y_{ji} + B''_i) - \\ - V_{j+1, i}(B'_i y_{j+1, i} + B''_i) = 0 \quad \begin{matrix} j = 2, 3, \dots, p_i - 1 \\ i = 0, 1, \dots, v. \end{matrix} \end{array} \right.$$

A (2) egyenletrendszer utolsó egyenletének jobb oldalán szereplő Q' mennyiség értelemszerűen extrakciónál a szolvens fázist, rektifikálásnál pedig a gőzfázist létrehozó folyamatot jellemzi. Ennek megfelelően értéke az általános extrakciónál az L_n fázissal egyensúlyba kerülő V_0 oldószerfázis előállítására szükséges oldószermennyiség:

$$(i) \quad Q' = V_0 B_0.$$

raffinátum refluxszal működő extraktornál az L_n — R raffinátum refluxnak oldószeres, azaz extrakt fázisbeli oldószertartalma, rektifikálásnál pedig

a visszaforrólóban közölt hőmennyiség. A rektifikáló teljes entalpiamérlegéből ekkor

$$(j) \quad Q' = D(H'_0 x_D + H''_0) + M(h'_v x_M + h''_v) - T h_T + L_0 \{(H'_0 - h'_0) x_D + (H''_0 - h''_0)\}.$$

(Erre a körülményre azért kell felhívni a figyelmet, mert rektifikálásnál $V_0 = 0$.)

Megjegyezzük, hogy a továbbiakban az általános alpművelet jelöléseinek megfelelően az oldószertartalomra vonatkozó egyenletekkel számolunk, de következtetéseink rektifikálás esetén is érvényben maradnak, csupán a B, b oldószertartalom helyett entalpiákat (H, h) kell írni.

Ha az (1) és (2) egyenletrendszeren az előző dolgozatunkban említett átalakítást végrehajtjuk (az l -edik egyenlet helyére az $l, l+1, \dots, n$ -edik egyenletek összegét írjuk), továbbá ha (c), (f), (g), (h) egyenleteket az (1) és (2) egyenletrendszerbe behelyettesítjük, átrendezés után a következő egyenletrendszereket kapjuk:

$$(3) \quad \begin{cases} (L_{0i} + G_i) K_i x_{1i} = L_{0i} x_{0i} + T x_T + V_0 y_0 - L_n x_n - K'_i G_i - K'_i L_{0i} \\ - L_{ji} x_{ji} + (L_{ji} + G_i) K_i x_{j+1,i} = T x_T + V_0 y_0 - L_n x_n - K'_i G_i - K'_i L_{0i} \end{cases} \quad \begin{matrix} j = 1, 2, \dots, p_i - 1 \\ i = 0, 1, \dots, v. \end{matrix}$$

$$(4) \quad \begin{cases} (L_{0i} + G_i) B'_i K_i x_{1i} = L_{0i} b'_i x_{0i} + T b_T + Q' - L_n (b'_v x_n + b''_v) - \\ - (B'_i K'_i + B''_i) G_i + (b''_i - B'_i K'_i - B''_i) L_{0i} \\ - L_{ji} b'_i x_{ji} + (L_{ji} + G_i) B'_i K_i x_{j+1,i} = T b_T + Q' - L_n (b'_v x_n + b''_v) - \\ - (B'_i K'_i + B''_i) G_i + (b''_i - B'_i K'_i - B''_i) L_{ji} \end{cases} \quad \begin{matrix} j = 1, 2, \dots, p_i - 1 \\ i = 0, 1, \dots, v, \end{matrix}$$

ahol értelemszerűen

$$(d') \quad G_i = \mathcal{E}_{ix} T + V_0 - L_n.$$

Itt ismeretlenek x_{ji} és L_{ji} értéke, a meghatározandó p_i fokozatszámot pedig a $T, x_T, x_0, x_n, x_k, y_0, L_0, V_0, b_T$ mennyiségek ismerete alapján fogjuk kiszámítani. A számításokhoz szükséges $K_i, K'_i, B'_i, B''_i, b'_i, b''_i$, illetve H'_i, H''_i, h'_i, h''_i ($i = 0, 1, \dots, v$) iránytangenseket és tengelymetszeteket a diagramról olvassuk le.

3. §. Az általános alpművelet elméleti fokozatszámának meghatározása

A (3) és (4) egyenletrendszerek a koncentrációkban lineárisak, azonban a koncentrációk L_{ji} együtthatói is ismeretlenek. Feladatunk — előző dolgozatunkhoz hasonlóan — az egyes egyenletrendszerekben az egyenletek számának meghatározása.

E feladat — éppen mivel az egyenletrendszer együtthatói is ismeretlenek — kétségtelenül nehezebb az előző dolgozatban tárgyalt feladatnál, megoldhatóságát azonban az biztosítja, hogy az egyenletrendszerek szerkezete, azaz az együtthatókból alkotott mátrixok igen egyszerűek. Különösképpen

amiatt, hogy — amint a (3) és (4) egyenletrendszer felírásán is látszik —, azok $v+1$ olyan csoportra oszlanak, amelyeken belül a húrpoligonokkal közelített függvények egyes szakaszait jellemző mennyiségek ($K_i, K'_i, B'_i, B''_i, b'_i, b''_i$) állandók. A gondolatmenet, amit követni fogunk, mármost a következő: az egyes fokozatcsoportokra felírt megfelelő (3) és (4) egyenletrendszerből a koncentrációk x_{ji} értékeit elimináljuk, s ezáltal az anyagáramok L_{ji} értékeire nem-lineáris egyenletrendszert nyerünk. Az így adódó egyenletrendszerből az L_{ji} értékekre tört-lineáris rekurziós formulához jutunk, amelynek felhasználásával, valamint figyelembe véve azt, hogy a lineáris szakaszok végpontjaihoz tartozó x_{0i} koncentráció-értékek ismeretesek, az egyes egyenlet-csoportok egyenleteinek p_i számát explicite meghatározhatjuk.

Tehát visszatérve a (3) és (4) egyenletrendszerre, azok egy-egy egyenlet-csoportját mátrix-egyenlet alakjában írjuk fel.

Az x_{ji} koncentrációk együtthatójával minden egyes egyenletet elosztva és bevezetve az

$$(5) \quad \begin{aligned} s_i &= \frac{1}{G_i} (\mathcal{E}_{ix} T x_T + V_0 y_0 - L_n x_n); \\ S_i &= \frac{1}{G_i} (\mathcal{E}_{ix} T b_T + Q' - L_n b_n) \quad (i = 0, 1, 2, \dots, v) \end{aligned}$$

jelöléseket, az $\mathbf{L}_i = [l_{jk}^{(i)}]$ p_i -edrendű nilpotens mátrixot, amelynek az

$$(6) \quad l_{j,j-1}^{(i)} = \frac{L_{j-1,i}}{(L_{j-1,i} + G_i) K_i} \quad \begin{aligned} j &= 2, 3, \dots, p_i \\ i &= 0, 1, 2, \dots, v \end{aligned}$$

elemein kívül valamennyi eleme 0, továbbá az

$$(7) \quad \begin{aligned} \mathbf{x}_i &= [x_{ji}]; \quad \mathbf{f}_i = [f_{ji}] = \left[\frac{1}{(L_{j-1,i} + G_i) K_i} \right]; \quad \mathbf{g}_i = [g_{ji}] = \left[\frac{L_{j-1,i}}{(L_{j-1,i} + G_i) K_i} \right]; \\ j &= 1, 2, \dots, p_i \\ i &= 0, 1, \dots, v \end{aligned}$$

vektorokat, valamint az \mathbf{e}_1 egységvektort, amelynek első eleme 1, a többi $p_i - 1$ eleme 0, ekkor a következő mátrixegyenleteket nyerjük:

$$(8) \quad (\mathbf{E} - \mathbf{L}_i) \mathbf{x}_i = \frac{L_{0i} x_{0i}}{(L_{0i} + G_i) K_i} \mathbf{e}_1 + G_i (s_i - K'_i) \mathbf{f}_i - K'_i \mathbf{g}_i, \quad (i = 0, 1, \dots, v)$$

$$(9) \quad \begin{aligned} &\left(\mathbf{E} - \frac{b'_i}{B'_i} \mathbf{L}_i \right) \mathbf{x}_i = \\ &= \frac{b'_i}{B'_i} \frac{L_{0i} x_{0i}}{(L_{0i} + G_i) K_i} \mathbf{e}_1 + \frac{G_i}{B'_i} \{S_i - (B'_i K'_i + B''_i)\} \mathbf{f}_i - \left(K'_i + \frac{B''_i - b''_i}{B'_i} \right) \mathbf{g}_i \\ &\quad (i = 0, 1, \dots, v). \end{aligned}$$

Ezekben az egyenletekben az ismeretlen L_{ji} értékek a koncentrációk együttható-mátrixain kívül az \mathbf{f}_i és \mathbf{g}_i vektorok komponenseiben fordulnak elő.

A kapott egyenletekből az is látható, hogy abban az esetben, ha az egyes fázisok oldószertartalmát meghatározó B , illetve b diagram (illetve a két entalpia-vonal) két párhuzamos egyenes, azaz $B'_i = b'_i$ (ill. $H'_i = h'_i$), akkor a (9) egyenlet a (8) egyenlettel azonos lesz. (Ennek kimutatásánál fel kell használni még azt is, hogy ebben az esetben a moláris lefolyás a táplálás felett, illetve alatt állandó. Lásd a ² lábjegyzetet.)

Ha most a (8) egyenletet balról $(\mathbf{E} - \mathbf{L}_i)^{-1}$ mátrixszal megszorozzuk és az \mathbf{x}_i vektorra így adódó kifejezést a (9) egyenletbe behelyettesítjük, akkor ezzel egyetlen olyan (p_i ismeretlenű) egyenletrendszert kapunk, amelyben az ismeretlenek csak az L_{ji} értékek. Ha figyelembe vesszük azt, hogy \mathbf{L}_i nilpotens mátrix és így

$$(\mathbf{E} - \mathbf{L}_i)^{-1} = \sum_{m=0}^{p_i-1} \mathbf{L}_i^m,$$

könnyű meggyőződni arról, hogy

$$\left(\mathbf{E} - \frac{b'_i}{B'_i} \mathbf{L}_i \right) (\mathbf{E} - \mathbf{L}_i)^{-1} = \frac{b'_i}{B'_i} \mathbf{E} + \frac{B'_i - b'_i}{B'_i} (\mathbf{E} - \mathbf{L}_i)^{-1}.$$

Ezt felhasználva, az L_{ji} értékekre tehát a következő egyenletrendszert kapjuk:

$$\begin{aligned} & \left\{ \frac{b'_i}{B'_i} \mathbf{E} + \frac{B'_i - b'_i}{B'_i} (\mathbf{E} - \mathbf{L}_i)^{-1} \right\} \left\{ \frac{L_{0i} x_{0i}}{(L_{0i} - G_i) K_i} \mathbf{e}_1 + G_i (s_i - K'_i) \mathbf{f}_i - K'_i \mathbf{g}_i \right\} = \\ & = \frac{b'_i}{B'_i} \frac{L_{0i} x_{0i}}{(L_{0i} + G_i) K_i} \mathbf{e}_1 + \frac{G_i}{B'_i} \{S_i - (B'_i K'_i + B''_i)\} \mathbf{f}_i - \left(K'_i + \frac{B''_i - b''_i}{B'_i} \right) \mathbf{g}_i. \end{aligned}$$

Ha beszorzás és rendezés után $(\mathbf{E} - \mathbf{L}_i)$ mátrixszal balról szorzunk és bevezetjük a

$$\begin{aligned} \overline{B}_i &= B'_i s_i + B''_i - S_i, \\ \overline{b}_i &= b'_i s_i + b''_i - S_i \end{aligned} \quad (10)$$

és

$$\Delta_{K'_i} = (B'_i - b'_i) K'_i + B''_i - b''_i \quad (11)$$

jelöléseket, némi átalakítás után a következő egyenletet nyerjük:

$$\begin{aligned} (B'_i - b'_i) \frac{L_{0i} x_{0i}}{(L_{0i} + G_i) K_i} \mathbf{e}_1 + G_i \overline{B}_i \mathbf{f}_i + (B''_i - b''_i) \mathbf{g}_i = \\ = \mathbf{L}_i \{ \Delta_{K'_i} \mathbf{g}_i + G_i (\overline{b}_i + \Delta_{K'_i}) \mathbf{f}_i \}. \end{aligned} \quad (12)$$

Tekintettel arra, hogy az \mathbf{L}_i nilpotens mátrixnak az első sorában minden elem 0, az \mathbf{e}_1 vektornak pedig az első kivételével minden eleme 0, külön választjuk a (12) egyenletrendszer első egyenletét a többitől:

$$(B'_i - b'_i) L_{0i} x_{0i} + G_i \overline{B}_i + (B''_i - b''_i) L_{0i} = 0. \quad (13)$$

Ez az egyenlet, mint látható, az ismert x_{0i} és az ismeretlen L_{0i} értékek között szolgáltat összefüggést.⁴

A (12) egyenletrendszer többi egyenletének segítségével, ha tekintetbe vesszük az \mathbf{L}_i mátrix (6) alakját, az L_{ji} értékekre rekurziós összefüggést írhatunk fel. Az \mathbf{L}_i mátrix elemei ugyanis megegyeznek a \mathbf{g}_i vektor elemeivel (elhagyva ennek első elemét). Mivel pedig az \mathbf{L}_i mátrix zérustól különböző elemei közvetlenül a főátló alatt foglalnak helyet, az $\mathbf{L}_i \mathbf{g}_i$ illetve $\mathbf{L}_i \mathbf{f}_i$ vektorokat a következő alakban nyerjük:

$$\left. \begin{aligned} \mathbf{L}_i \mathbf{g}_i &= [g_{ji} g_{j-1,i}], & (g_{-1,i} &= 0) \\ \mathbf{L}_i \mathbf{f}_i &= [g_{ji} f_{j-1,i}], & (f_{-1,i} &= 0) \end{aligned} \right\} \begin{array}{l} j=0, 1, 2, \dots, p_i-1 \\ i=0, 1, 2, \dots, v \end{array}$$

A (12) egyenletrendszer többi egyenlete tehát a következő:

$$G_i \bar{B}_i f_{ji} + (B_i'' - b_i'') g_{ji} = g_{ji} \{ \Delta_{K_i'} g_{j-1,i} + G_i (\bar{b}_i + \Delta_{K_i'}) f_{j-1,i} \}, \quad j=1, 2, \dots, p_i-1.$$

Az f_{ji} és g_{ji} mennyiségek helyére (7) alapján az L_{ji} értékeket behelyettesítve, a következő összefüggést kapjuk:

$$\frac{G_i \bar{B}_i}{L_{ji}} + B_i'' - b_i'' = \frac{\Delta_{K_i'} L_{j-1,i} + G_i (\bar{b}_i + \Delta_{K_i'})}{(L_{j-1,i} + G_i) K_i}, \quad j=1, 2, \dots, p_i-1.$$

Összevonás és az

$$(14) \quad A_i = \Delta_{K_i'} - K_i (B_i'' - b_i'')$$

jelölés bevezetése után L_{ji} meghatározására az alábbi tört-lineáris rekurziós formulát nyerjük:

$$(15) \quad \frac{1}{\frac{L_{ji}}{G_i \bar{B}_i}} = \frac{A_i \frac{L_{j-1,i}}{G_i \bar{B}_i} + \frac{1}{\bar{B}_i} (\bar{b}_i + A_i)}{K_i \frac{L_{j-1,i}}{G_i \bar{B}_i} + \frac{K_i}{\bar{B}_i}}.$$

Tekintettel arra, hogy a lineáris rekurziós formulák megoldására egyszerű explicit összefüggés van, célszerűnek látszik a kapott (15) tört-lineáris rekurziós formulát transzformációval linearizálni. Könnyen belátható, hogy ez egyszerű eltolási transzformációval elérhető. Vezessük be ugyanis

$$\frac{1}{\frac{L_{ji}}{G_i \bar{B}_i}} \quad \text{helyett az} \quad \frac{1}{\frac{L_{ji}}{G_i \bar{B}_i} + w}$$

⁴ (13) az i -edik és n -edik fokozatok közötti oszloprészre vonatkozó anyag- és oldószert (entalpia) egyenleg, azaz a (3) és (4) egyenletrendszer egymásnak megfelelő egy-egy egyenletének egybevetéséből is kiadódik.

kifejezést. Ekkor a (15) összefüggés némi átalakítás után a következő alakra hozható:

$$\frac{1}{\frac{L_{ji}}{G_i \bar{B}_i} + w} = \frac{A_i}{K_i + w A_i} + \frac{\frac{\bar{b}_i}{\bar{B}_i} K_i}{(K_i + w A_i)^2} \cdot \frac{1}{\frac{L_{j-1,i}}{G_i \bar{B}_i} + \frac{K_i + w(\bar{b}_i + A_i)}{\bar{B}_i(K_i + w A_i)}}.$$

Ha most w értékét úgy határozzuk meg, hogy kielégítse a

$$(16) \quad w = \frac{K_i + w(\bar{b}_i + A_i)}{\bar{B}_i(K_i + w A_i)}$$

egyenletet, akkor az

$$\frac{1}{\frac{L_{ji}}{G_i \bar{B}_i} + w}$$

kifejezésre már lineáris rekurziós formulát kapunk:

$$(17) \quad \frac{1}{\frac{L_{ji}}{G_i \bar{B}_i} + w} = \frac{A_i}{K_i + w A_i} + \frac{\frac{\bar{b}_i}{\bar{B}_i} K_i}{(K_i + w A_i)^2} \cdot \frac{1}{\frac{L_{j-1,i}}{G_i \bar{B}_i} + w}.$$

A (17) egyenletből w értékére.

$$(18) \quad w_{\pm} = \frac{1}{2 A_i \bar{B}_i} \left\{ \bar{b}_i - K_i \bar{B}_i + A_i \pm \sqrt{(\bar{b}_i - K_i \bar{B}_i + A_i)^2 + 4 A_i K_i \bar{B}_i} \right\}$$

adódik. Itt és a továbbiakban az áttekinthetőség kedvéért w -nek, majd az ebből származó u és v -nek i -től való függését nem tüntetjük fel. A (17) formula állandó együtthatójú lineáris elsőrendű differenciaegyenlet, amelynek megoldásához egyszerű mátrixelméleti megfontolással juthatunk. Ha ugyanis a differenciaegyenletet algebrai egyenletrendszerként írjuk fel és bevezetjük a

$$(19) \quad \begin{aligned} \frac{\bar{b}_i K_i}{\bar{B}_i(K_i + w_{\pm} A_i)^2} &= v_{\pm}, \\ \frac{A_i}{K_i + w_{\pm} A_i} &= u_{\pm} (v_{\pm} - 1), \\ z_j &= \frac{1}{\frac{L_{ji}}{G_i \bar{B}_i} + w_{\pm}} \end{aligned}$$

jelöléseket, amelyekkel a (17) egyenletrendszer (a \pm indexet ideiglenesen elhagyva)

$$(17') \quad z_j = u(v - 1) + v z_{j-1}$$

alakban írható fel, továbbá a $(p_i - 1)$ -edrendű \mathbf{E} egységmátrixot, az $\mathbf{N} = [n_{\alpha\beta}]$, $(p_i - 1)$ -edrendű nilpotens mátrixot, amelynek $n_{\alpha, \alpha-1} = 1$ ($\alpha = 2, 3, \dots, p_i - 1$) elemein kívül valamennyi eleme 0, valamint a $(p_i - 1)$ -edrendű \mathbf{e}_1 egységvektort, a csupa 1 elemet tartalmazó \mathbf{e} vektort és a $\mathbf{z} = [z_j]$ ($j = 1, 2, \dots, p_i - 1$) vektort, akkor a (17') egyenletrendszer a következő lesz:

$$(\mathbf{E} - v\mathbf{N})\mathbf{z} = vz_0 \mathbf{e}_1 + u(v - 1) \mathbf{e}.$$

Innen, mivel

$$(\mathbf{E} - v\mathbf{N})^{-1} = \sum_{m=0}^{p_i-2} (v\mathbf{N})^m = \begin{bmatrix} 1 & & & & & \\ v & 1 & & & & \\ v^2 & v & 1 & & & \\ . & v^2 & v & 1 & & \\ . & . & . & . & . & \\ . & . & . & . & . & \\ . & . & . & . & . & \\ & & & & & 1 \\ v^{p_i-2} & v^{p_i-3} & & & & v & 1 \end{bmatrix}$$

a keresett z_j elemekre

$$(20) \quad z_j = (u + z_0) v^j - u, \quad (j = 1, 2, \dots, p_i - 1)$$

adódik. A (19) jelöléseket visszahelyettesítve, látható hogy L_{ji} ismeretlen értékei L_{0i} függvényében explicit alakban nyerhetők, L_{0i} pedig a (13) egyenlet alapján x_{0i} ismeretében könnyen kiszámítható. Bennünket azonban elsősorban az egyes fokozatcsoportokban a *fokozatok száma*, azaz p_i értéke érdekel. Vegyük tekintetbe, hogy bármelyik fokozatcsoport első egyenlete az előző csoport utolsó egyenleteként is felírható, hiszen maguk az osztáspontok, amelyekkel az egyensúlyi görbét és a többi diagramot lineáris szakaszokra osztottuk, egyaránt tekinthetők akár a tőlük balra, akár a tőlük jobbra levő szakasz pontjaiként. Ha tehát ily módon a (17) egyenletrendszert egy p_i -edik egyenlettel kiegészítjük, amelyben $L_{p_i i} = L_{0, i+1}$ (ez utóbbi viszont a (13) egyenletből $x_{0, i+1}$ segítségével kiszámítható), akkor az egyenletrendszer (20) megoldása egy p_i -edik összefüggéssel bővül:

$$(21) \quad z_{p_i} = (u + z_0) v^{p_i} - u.$$

Tekintettel arra, hogy z_{p_i} egyszerű összefüggésben van $L_{p_i i}$ értékével, a (21) egyenletből p_i közvetlenül kiszámítható:

$$(22) \quad p_i = \frac{\log \frac{u + z_{p_i}}{u + z_0}}{\log v}.$$

Ha most a (13) egyenletből L_{0i} értékét kifejezzük és behelyettesítjük z_{p_i} és z_0 képletébe, akkor

$$(23) \quad A_i = (B'_i - b'_i) x_{0i} + (B''_i - b''_i)$$

jelöléssel

$$z_0 = \frac{1}{w_{\pm} - \frac{1}{A_i}}$$

és

$$z_{p_i} = \frac{1}{w_{\pm} - \frac{1}{A_{i+1}}}$$

adódik. Ezt behelyettesítve a (22) formulába, egyszerű átalakítás után

$$(24) \quad p_i = \frac{\log \left\{ \left(w_{\pm} + \frac{1}{u_{\pm}} \right) A_{i+1} - 1 \right\} (w_{\pm} A_i - 1)}{\left\{ \left(w_{\pm} + \frac{1}{u_{\pm}} \right) A_i - 1 \right\} (w_{\pm} A_{i+1} - 1)} \log v_{\pm}$$

adódik. Ha végül figyelembe vesszük azt, hogy w_{\pm} (18) alatti kifejezéséből

$$K_i + w_{\pm} A_i = \frac{1}{2 \bar{B}_i} \left\{ \bar{b}_i + K_i \bar{B}_i + A_i \pm \sqrt{(\bar{b}_i + K_i \bar{B}_i + A_i)^2 - 4 K_i \bar{B}_i \bar{b}_i} \right\}$$

és

$$\frac{1}{K_i + w_{\pm} A_i} = \frac{1}{2 K_i \bar{b}_i} \left\{ \bar{b}_i + K_i \bar{B}_i + A_i \mp \sqrt{(\bar{b}_i + K_i \bar{B}_i + A_i)^2 - 4 K_i \bar{B}_i \bar{b}_i} \right\},$$

tehát

$$(25) \quad \frac{1}{K_i + w_{\pm} A_i} = \frac{\bar{B}_i}{K_i \bar{b}_i} (K_i + w_{\mp} A_i),$$

akkor (19) megfelelő összefüggésébe behelyettesítve

$$(26) \quad v_{\pm} = \frac{\bar{b}_i + K_i \bar{B}_i + A_i \mp \sqrt{(\bar{b}_i + K_i \bar{B}_i + A_i)^2 - 4 K_i \bar{B}_i \bar{b}_i}}{\bar{b}_i + K_i \bar{B}_i + A_i \pm \sqrt{(\bar{b}_i + K_i \bar{B}_i + A_i)^2 - 4 K_i \bar{B}_i \bar{b}_i}},$$

és mivel ugyancsak (19)-ből

$$\frac{1}{u_{\pm}} = \frac{\frac{\bar{b}_i}{\bar{B}_i} \frac{K_i}{(K_i + w_{\pm} A_i)^2} - 1}{\frac{A_i}{K_i + w_{\pm} A_i}},$$

a (25)-ös összefüggés felhasználásával

$$\frac{1}{u_{\pm}} = w_{\mp} - w_{\pm}$$

adódik. Ezt behelyettesítve p_i (24) kifejezésébe, végül azt kapjuk, hogy

$$(27) \quad p_i = \frac{\log \frac{(w_{\pm} \Delta_i - 1)(w_{\pm} \Delta_{i+1} - 1)}{(w_{\pm} \Delta_{i+1} - 1)(w_{\mp} \Delta_i - 1)}}{\log v_{\pm}}.$$

Az itt szereplő v_{\pm} és w_{\pm} valamint Δ_i jelentését a (26), (18), illetve (23) formula szolgáltatja, továbbá az ezekben szereplő további jelöléseket megadják a (11), (10) és (5) formulák. A kapott (27) eredményből annak egyértelműsége is kiolvasható. Ugyanis a v_{\pm} és w_{\pm} kifejezéseiben szereplő gyököket akár pozitív, akár negatív előjellel véve, ugyanarra az eredményre jutunk.

A keresett fokozatszámot a táplálásig terjedő oszlopprészben

$$(28) \quad k = \sum_{i=0}^{x-1} p_i$$

a táplálás után oszlopprészben pedig

$$(28') \quad n - k = \sum_{i=x}^r p_i$$

adja.

4. §. Az általános alaplóműveletre vonatkozó eredmény alkalmazása extraktorra és rektifikálóra

1. Közepső táplálás és reflux nélkül működő extraktor

Feltételek:

$$T = 0; \quad Q' = V_0(B'_v y_0 + B''_v).$$

Tehát a számításhoz szükséges képletek a következők:

$$s = \frac{V_0 y_0 - L_n x_n}{V_0 - L_n}; \quad S = \frac{V_0(B'_v y_0 + B''_v) - L_n(b'_v x_n + b''_v)}{V_0 - L_n},$$

A_i , \bar{b}_i , \bar{B}_i számértékeit pedig az

$$(14) \quad A_i = (B'_i - b'_i) K'_i + (1 - K_i) (B''_i - b''_i)$$

$$(10) \quad \bar{b}_i = b'_i s + b''_i - S$$

$$\bar{B}_i = B'_i s + B''_i - S$$

képletekkel nyerjük.

2. Raffinátum refluxszal és extrakt-refluxszal működő extraktor

Feltételek:

$$x_n = y_0 = x_R; \quad V_0 = L_n - R;$$

$$x_0 = y_{10} = x_E; \quad V_{10} = L_0 + E;$$

$$Q' = V_0(B'_v x_R + B''_v).$$

Tekintettel arra, hogy L_n értéke ismeretlen, V_0 is ismeretlen. Így a fenti kifejezések közvetlen helyettesítése nem alkalmazható. Felhasználva azonban az egész készüléknek a B anyagra vonatkozó mérlegét, a (29) összefüggést, valamint s és S definícióit,

$$s = \frac{\mathcal{E}_{ix} T x_T + (L_n - R) x_R - L_n x_R}{\mathcal{E}_{ix} T + (L_n - R) - L_n} = \begin{cases} x_E & \text{a táplálás felett,} \\ x_R & \text{a táplálás alatt.} \end{cases}$$

$$\begin{aligned} S &= \frac{\mathcal{E}_{ix} T b_T + V_{10}(B'_0 x_E + B''_0)}{\mathcal{E}_{ix} T + (L_n - R) - L_n} + \\ &+ \frac{L_n(b'_v x_R + b''_v) - T b_T - L_0(b'_0 x_E + b''_0) - L_n(b'_v x_R + b''_v)}{\mathcal{E}_{ix} T + (L_n - R) - L_n} = \\ &= \frac{(\mathcal{E}_{ix} - 1) T b_T + (L_0 - E) \Delta_{x_E} + E(b'_0 x_E + b''_0)}{T - R}, \end{aligned}$$

A_i , \bar{b}_i , B_i értékeit most is a (14) és (10) képletekből nyerjük.

3. Rektifikáló

Feltételek:

$$x_n = x_M; \quad V_0 = D; \quad L_n = M;$$

$$x_0 = y_{10} = x_D; \quad V_{10} = L_0 + D.$$

A visszaforralóban közölt Q' hőmennyiség a 285. oldalon található (j) képlet alapján:

$$Q' = (L_0 + D) \Delta_{x_D} + D(h'_0 x_D + h''_0) + M(h'_v x_M + h''_v) - T h_T.$$

Ezzel s és S értékei teljes összhangban a refluxos extraktor fenti kifejezéseivel

$$s = \begin{cases} x_D & \text{a táplálás felett,} \\ x_M & \text{a táplálás alatt,} \end{cases}$$

$$S = \frac{(\mathcal{E}_{ix} - 1) T h_T + (L_0 + D) \Delta_{x_D} + D(h'_0 x_D + h''_0)}{\mathcal{E}_{ix} T - M}$$

alakban adódnak, A_i , \bar{b}_i , B_i értékeit ismét a (14) és (10) képletek adják meg.

4. Teljes refluxszal működő rektifikáló

Az előbbi specifikáló feltételeket még azzal kell kiegészíteni, hogy

$$T = D = M = 0.$$

Ebből még az következik, hogy $L_i = V_{i+1}$, ami annyit jelent, hogy teljes reflux alkalmazása mellett a munkavonal L és V változó volta ellenére is az $y(x)$ koordinátarendszer átlójával esik egybe. Így a teljes refluxszal működő rektifikáló elméleti fokozatszám az entalpiavonalak alakjától függetlenül előző dolgozatunk 4. paragrafus 3. pontjában közölt képlettel számítható ki. Megjegyezzük végül, hogy teljes reflux esetén $S = \infty$. Ennek egyszerű helyettesítése a (14) és (10) képletekbe nem lehetséges. Viszont S -el való osztás után a (18) és (26) képletekben a határátmenet $S \rightarrow \infty$ esetén elvégezhető és az így nyert elméleti fokozatszám előző dolgozatunk fent említett képletével megegyezik.

*

A specifikáló feltételekkel nyert \bar{b}_i és \bar{B}_i képleteket w_{\pm} (18)-as és v_{\pm} (26)-os képleteibe kell behelyettesíteni, majd ezek felhasználásával (27)-ből p_i értékét ki lehet számítani. Ezeket a helyettesítéseket azonban — bonyolult és áttekinthetetlen formulák elkerülése végett — célszerű konkrét esetekben numerikusan végrehajtani.

Jelölések

A	az értékes anyagot hordozó oldószer
B	az extraháló oldószer
C	az értékes anyag
$A_{\text{mól}}, B_{\text{mól}}, C_{\text{mól}}$	az A , B , C anyag mennyisége mólban
B	a B anyag oldószermentes alapú mólviszonya a V fázisban
b	a B anyag oldószermentes alapú mólviszonya az L fázisban
b_T	a B anyag oldószermentes alapú mólviszonya a T táplálékban
D	a desztillátum moláris sebessége
E	az extraktum moláris sebessége
H	a V fázis moláris entalpiája
h	az L fázis moláris entalpiája
k	a táplálást magába fogadó fokozat sorszáma
L	az értékes anyagban szegényedő fázis moláris sebessége ($A + C$ -vel kifejezve)
M	a desztillációs maradék moláris sebessége
n	az elméleti fokozatok száma
Q'	V fázis előállítására szükséges B anyag mennyisége, illetve rektifikálásnál a visszaforrálóban közölt hőmennyiség
q	a táplálásnak L -hez csatlakozó hányada
R	az extraktorból kilépő raffinátum moláris sebessége
T	a táplálás moláris sebessége
V	az értékes anyagban dúsuló fázis moláris sebessége ($A + C$ -vel kifejezve)
x	az értékes anyag móltörtje L -ben
y	az értékes anyag móltörtje V -ben

$$\mathcal{E}_{lk} = \begin{cases} 1, & \text{ha } k-l \geq 0 \\ 0, & \text{ha } k-l < 0 \end{cases} \quad \text{az egységugrás függvénye}$$

E egységmátrix
e₁ első egységvektor

A dolgozatban szereplő közvetett jelölések jegyzéke

A_i	288. oldal	(14)
$\overline{B_i}, \overline{b_i}$	287. „	(10)
Δ_i	291. „	(23)
$\Delta_{K'}$	287. „	(11)
G_i	285. „	(d')
p_i	292. „	(27)
q	282. „	(e)
S_i, s_i	286. „	(5)
v	291. „	(26)
w	289. „	(18)
K_i, K'_i	282. „	(f)
B'_i, B''_i	282. „	(g)
b'_i, b''_i	282. „	(h)

(Beérkezett: 1959. május 15.)

IRODALOM

- [1] JUNG G.—RÓZSA P.—SÁRKÁNY GY.: „Ellenáramú szétválasztó vegyipari alpműveletek elméleti fokozatszámának meghatározásáról, I.” *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* **2** (1957) 227—245.
- [2] ROBINSON, C. S.—GILLILAND, E. R.: *Elements of fractional distillation*. McGraw-Hill, New York, 1950. (4. ed.)
- [3] KIRSCHBAUM, E.: *Destillier- und Rektifiziertchnik*. Springer, Berlin, 1950. (2. Aufl.)
- [4] THUM, O.: „Direkte Berechnung der theoretischen Bodenzahl bei Destillation idealer Zweistoffgemische mit unterschiedlichen Verdampfungswärmen und bei endlichem Rücklaufverhältnis.” *Chemie-Ingenieur-Technik* **29** (1957) 675—678.

DETERMINATION OF THE THEORETICAL STAGE-NUMBER OF COUNTER-CURRENT DISTRIBUTIVE CHEMICAL UNIT OPERATIONS, II.

by

P. RÓZSA and GY. SÁRKÁNY

In this paper, as has been done in [1], an attempt is made to give an explicit formula for the computation of the theoretical stage-number of extractors with two feed solutions and without reflux, using such an extractor as a general unit operation model.

In previous work the supposition was made that the mass velocity of the two phases in countercurrent flow can be characterised by a velocity-like constant parameter, which can be computed. In this work this supposition is not maintained, consequently a more general case is investigated in which, besides the concentrations in the single stages, also the variable mass velocities (L_i and V_i) of the stages are unknown.

In the first chapter the basic relations are given (material balance resp. heat balance, and equilibrium relations between concentrations of the phases) which were, in part, used in our previous communication, and which are, in part, introduced here the first time. These relations are used in the next chapter as a starting point for obtaining those equations which serve for the determination of the stage-number. In the setting up of these equations here too the practice is followed of approximating by chord-polygons the curves which represent empirical relations. In order to determine the stage-number the systems of equations are divided into groups according to chord-polygons, and the number of the equations within the single groups determined. The mathematical manipulation of the problem leads to a system of non-linear equations; for the unknowns fractional linear recursion formulae are found. With the solution of the latter the problem proper, i. e. the determination of the number of equations, is amenable to solution. The results i. e. the equations suitable to serve for computations are to be found on pp. 292 (equations (27), (28), (28').) The notations used in these equations are explained at the end of this paper. Finally, the specifying conditions for extractors and rectifiers without mid-point feed inlet are given; by the help of these specifications the generalized result can be applied also to these operations.

ОПРЕДЕЛЕНИЕ ЧИСЛА ТЕОРЕТИЧЕСКИХ СТУПЕНЕЙ ПРОТИВОТОЧНЫХ РАЗДЕЛЯЮЩИХ ПРОЦЕССОВ ХИМИЧЕСКОЙ ТЕХНОЛОГИИ, II.

P. RÓZSA и GY. SÁRKÁNY

В настоящей работе, — как в предыдущем сообщении [1], — дается расчетная формула для определения числа теоретических ступеней питаемого в середине экстрактора без возврата (являющегося общей моделью) с применением некоторых приближенных предположений. В предыдущей работе предполагалось, что расходы фаз являются постоянными и характеризуются рассчитываемыми параметрами (типа скорости). В настоящем предположении это не предполагается, значит исследуем тот более общий случай, когда в отдельных ступенях кроме концентрации меняющиеся расходы фаз (L_i и V_i) являются также неизвестными.

В первой главе находятся те основные зависимости (материальные и тепловые балансы, в зависимости фазового равновесия), использованные в частности в предыдущей работе, на основе которых составляются в следующей главе системы уравнений, необходимые для определения числа ступеней. При этом кривые, полученные экспериментальным путем, замещаются полигонами. Число ступеней определяется таким образом, что системы уравнений разделяются на группы, выбранные соответственно полигонам, и потом

определяются числа уравнений отдельных групп. Задача ведет к нелинейной системе уравнений, а для неизвестных получается дробно-линейная рекурзивная формула. С решением последней можно решить и собственную задачу — определить число уравнений. Полученные уравнения, пригодные для расчета, находятся на страницах 292 [(27), (28), (28')]. Обозначения в этих формулах можно найти в конце работы. В заключении даются те специфицирующие условия, при помощи которых общий результат можно применять в случае экстрактора (без питания в середине, экстрактора с возвратом и ректификационной колонны).

TÖBBSZABADSÁGFOKÚ LENGŐRENDSZER GERJESZTETT LENGÉSEI VIZSGÁLATÁNAK EGY MÓDSZERE S NÉHÁNY ALKALMAZÁSA

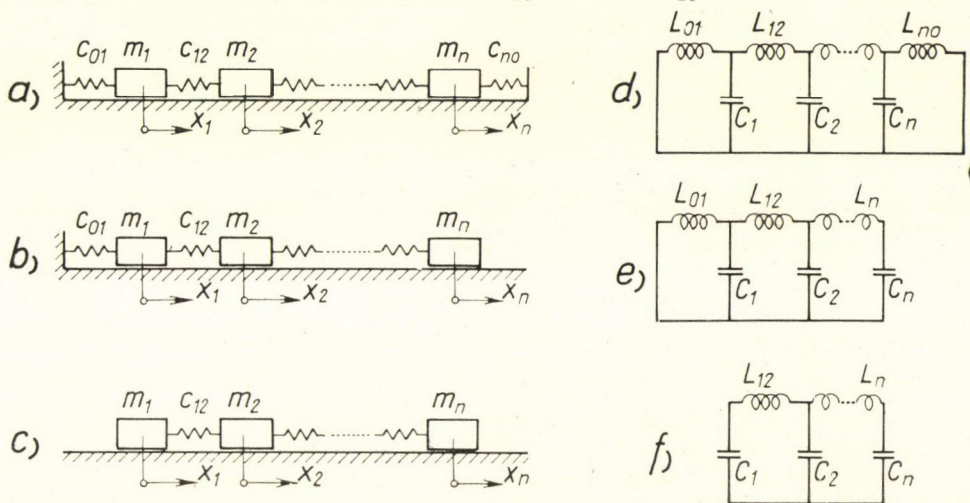
BOSZNAY ÁDÁM

1. Bevezetés, célkitűzés

Véges sok szabadságfokú lengőrendszer periódusos gerjesztő erők hatására kialakuló, a gerjesztő hatás periódusával megegyező periódusú lengése amplitúdójának meghatározása elméleti és gyakorlati szempontból is jelentős feladat, s mind a műszaki és elméleti mechanikában, mind a műszaki és elméleti villamoságtanban gyakran találkozunk vele. Ezt a feladatot kell például megoldani belsőégésű motor főtengelye csavaró lengés okozta többletigénybevételeinek meghatározásakor, vagy pl. mechanikai vagy villamos szűrők vizsgálatánál. Dolgozatunk célja: ennek a feladatnak a megoldására alkalmas, előnyösnek ítélt eljárás és az eljárás néhány alkalmazási lehetőségének bemutatása. A dolgozat kifejezés módja elsősorban a mechanikai lengőrendszereket tartja szem előtt, de minden eredménye átvihető az ismert módon megfelelő villamos kapcsolásokra is.

2. Az eljárás ismertetése

A szóbanforgó eljárás közvetlenül és a legegyszerűbben az 1. ábra szerinti mechanikai lengőrendszer-modellekre vagy a velük egyenértékű és az ábrán



1. ábra

ugyancsak feltüntetett villamos kapcsolásokra alkalmazható. További vizsgálatainkban az 1b ábrán rajzolt modellt tartjuk szem előtt, az ezzel kapcsolatban nyert eredmények jól ismert módon átvihetők az 1. ábra többi modelljére.

Az 1b ábrán látható modell kis kitérésű gerjesztett lengéseit az

$$\mathbf{E} = \begin{bmatrix} m_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & m_n \end{bmatrix};$$

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & 0 & \dots & 0 \\ u_{12} & u_{22} & u_{23} & 0 & \dots & 0 \\ 0 & u_{23} & u_{33} & u_{34} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & u_{n-1,n} & u_{nn} \end{bmatrix} =$$

$$= \begin{bmatrix} \underbrace{1} & \underbrace{2} & \underbrace{k-1} & k & \underbrace{k+1} & \underbrace{n-1} & \underbrace{n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \quad (1)$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{és} \quad \mathbf{p} = \begin{bmatrix} P_1(t) \\ \vdots \\ P_n(t) \end{bmatrix}$$

mátrixok bevezetésével az

$$(1) \quad \mathbf{E} \cdot \ddot{\mathbf{x}} = -\mathbf{U} \cdot \mathbf{x} + \mathbf{p}$$

mátrixdifferenciálegyenlet írja le. Periódusos $P_k(t)$, ($k = 1, \dots, n$) gerjesztő erők esetén (1) lineáris felépítése miatt érvényes a szuperpozíció elve, vagyis a feladat megoldását

$$\mathbf{p}_0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ P_{0i} \cdot \sin \omega t \\ \vdots \\ 0 \end{bmatrix}$$

alakú gerjesztő erőkhöz tartozó megoldások összegeként nyerjük. A gerjesztő hatás periódusával megegyező vagy röviden állandósult lengésnek

$$\mathbf{x} = \begin{bmatrix} x_{01} \\ \vdots \\ x_{0n} \end{bmatrix} \cdot \sin \omega t = \mathbf{x}_0 \cdot \sin \omega t$$

megoldásvektor felel meg. (1)-be visszahelyettesítve és $\sin \omega t$ -vel egyszerűsítve

$$(2) \quad (-\omega^2 \mathbf{E} + \mathbf{U}) \cdot \mathbf{x}_0 = \mathbf{p}_0$$

inhomogén lineáris algebrai egyenletrendszert kapjuk \mathbf{x}_0 meghatározására
Részletesen kiírva:

$$(3) \left\{ \begin{array}{l} (-\omega^2 m_1 + u_{11}) x_{01} + u_{12} \cdot x_{02} = 0 \\ u_{12} \cdot x_{01} + (-\omega^2 m_2 + u_{22}) x_{02} + u_{23} \cdot x_{03} = 0 \\ \dots\dots\dots \\ \dots\dots u_{i-1,i} \cdot x_{0,i-1} + (-\omega^2 m_i + u_{ii}) x_{0i} + u_{i,i+1} \cdot x_{0,i+1} \dots\dots = P_{0i} \\ \dots\dots\dots \\ u_{n-1,n} \cdot x_{0,n-1} + (-\omega^2 m_n + u_{nn}) x_{0n} = 0 . \end{array} \right.$$

Módszerünk alapgondolata az, hogy \mathbf{x}_0 meghatározását $\alpha = \omega$ kör-frekvenciájú saját lengés lengésképeinek meghatározására vezetjük vissza. Ez a gondolat a (3) egyenletrendszerrel kapcsolatban úgy szemléltethető, hogy az i -edik egyenlet jobboldalát az alábbi módon a baloldalra vesszük át (feltéve, hogy $x_{0i} = 0$):

$$u_{i-1,i} \cdot x_{0,i-1} + \left[-\omega^2 \left(m_i + \frac{P_{0i}}{\omega^2 \cdot x_{0i}} \right) + u_{ii} \right] x_{0i} + u_{i,i+1} \cdot x_{0,i+1} = 0.$$

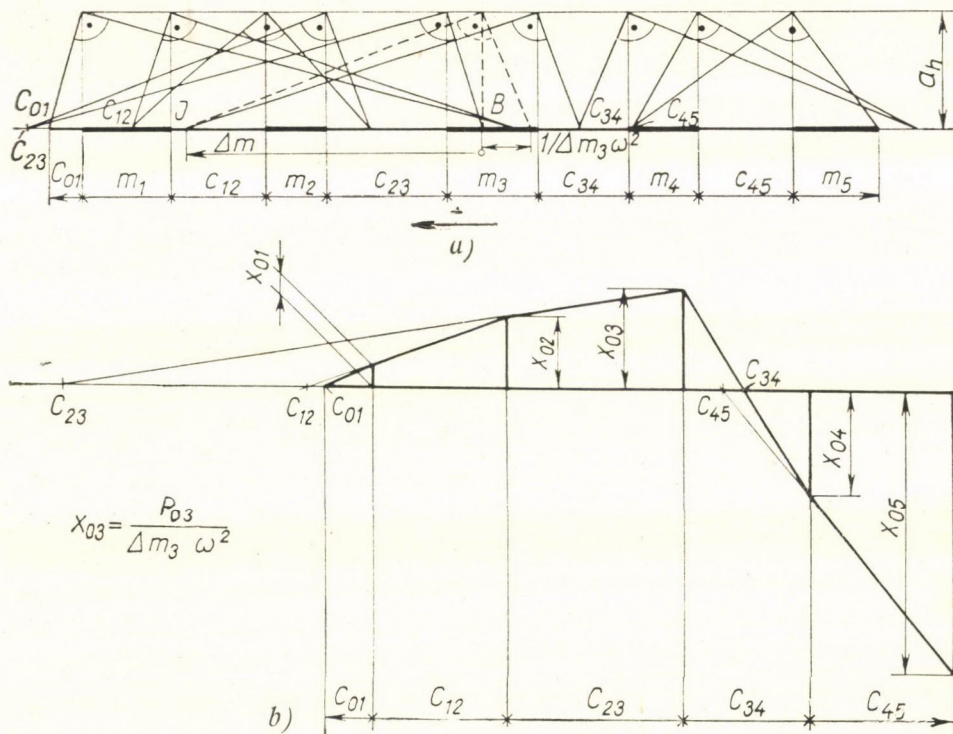
Ez után az átalakítás után (3) alakja olyan, mint az eredetitől csak az i -edik tömeg tekintetében különböző lengőrendszer ω körfrekvenciájú sajátlengésformáját meghatározó egyenletrendszer; ω tehát a sajátkörfrekvencia szerepét tölti be. A módosított i -edik tömeg:

$$(4) \dots \quad m'_i = m_i + \frac{P_{0i}}{\omega^2 \cdot x_{0i}}.$$

A fent x_{0_k} -kra felírt homogén lineáris egyenletrendszer együtthatói most mind ismeretesek, kivéve x_{0_i} együtthatóját az i -edik egyenletben, ebben ugyanis maga az x_{0_i} ismeretlen is szerepel. Az a feltétel, hogy az egyenletrendszernek legyen a triviálisól különböző megoldása, x_{0_i} számára ad meghatározó egyenletet. x_{0_i} -ből m'_i számítható, s így már minden együttható ismeretes.

Az egyenletrendszer számításhoz megoldására az alábbi út kínálkozik. Bár az ismert x_{0i} már egyértelműen meghatározza az összes ismeretlent,

vegyük fel mégis egyelőre x_{01} -et tetszés szerint, legyen ez a tetszés szerint felvett érték x_{01}^* . Ebből — sorról sorra haladva — a többi, ehhez az x_{01}^* értékhez tartozó s ugyancsak csillaggal jelölt ismeretlen is meghatározható. Az $(i - 1)$ -edik sorból adódó x_{0i}^* általában nem egyezik meg az előbb számított (helyes) x_{0i} -vel. Az eddig kiszámított, csillaggal jelölt mennyiségekből a (3) egyenletrendszert kielégítő értékeket azoknak x_{0i}/x_{0i}^* -al való szorzásával kapjuk. Ezután sorról sorra továbbhaladva meghatározhatjuk a többi ismeretlent. Az utolsó egyenlet kontrollként szolgál. A fent vázolt — csillagos mennyiségekkel történő — számítás az egyenletrendszer mátrixának kontinuáns voltát kihasználva elkerüli a determinánsokkal való számolást.



2. ábra

$i = 1$ vagy $i = n$ esetén (azaz ha éppen az első vagy az utolsó tömegre működtetünk gerjesztést) az eljárás annyiban egyszerűbb, hogy a csillagos mennyiségekkel való számolás elmarad.

Az eljárás *grafikus változata* az eredetileg sajátkörfrekvencia és sajátlengésforma meghatározására konstruált Kohn-féle szerkesztést [2] alkalmazza.¹ A szerkesztést azzal készítjük elő, hogy alkalmas m_x [cm/kpsec²cm⁻¹] tömeg-

¹ Sajátkörfrekvenciák és sajátlengésformák meghatározására szolgáló eljárás gerjesztett lengések vizsgálatára való alkalmazásának gondolatát először HOLZERNÉL találjuk meg [9]. A Kohn-szerkesztést tudomása szerint a szerző alkalmazta először erre a célra [3], majd SÁLYI kiterjesztette alkalmazhatóságát sebességgel arányos csillapítással kiegészített modell gerjesztett lengésének vizsgálatára [4].

és $c_x [\text{cm/cmkp}^{-1}]$ rugóállandómérték felvétele után felrajzoljuk a lengőrendszer úgynevezett vázát (2a ábra), majd tőle $a_h = \frac{\sqrt{m_x \cdot c_x}}{\omega}$ távolságban vele párhuzamosat húzunk. Ábránkon példaképpen ötszabadságfokú, egyik végén kötött lengőrendszer vázát rajzoltuk fel, s a 3. tömeget támadja $P_{03} \cdot \sin \omega t$ gerjesztőerő.

A szerkesztés abban áll, hogy a modellváznak mind a bal-, mind a jobboldali végpontjából kiindulva megrajzoljuk az ábra alapján követhető módon a Kohn-szerkesztés törtvonalát, de mindkét oldalról csak a gerjesztett tömeg képéig, ábránkon a B, illetve J pontig. Ha ez a két pont egybeesne, az [2] szerint azt jelentené, hogy ω a lengőrendszer egyik sajátkörfrekvenciája volna, azaz a rezonancia esete állna fenn. Példánkban a \vec{BJ} — tömegmértékben Δm_3 — B-től J felé irányított távolság zérustól különböző és a berajzolt pozitív iránynyal megegyező. Ez azt jelenti, hogy m_3 helyett $m'_3 = m_3 + \Delta m_3$ -at véve, a módosított lengőrendszernek ω az egyik sajátkörfrekvenciája lesz. A módosított lengőrendszer tehát külső erőhatás nélkül ω körfrekvenciájú sajátlengést végezhetne; a módosított tömeg $x_3 = x_{03} \cdot \sin \omega t$ törvényszerűség szerint lenghetne. Ennek a módosított rendszernek ezt a lengésképét valamelyik amplitúdó tetszés szerinti felvétele után a $C_{01}, C_{12}, C_{23}, C_{34}, C_{45}$ (részben valós, részben képzetes) csomópontok ismeretében könnyen megrajzolhatjuk. $\Delta m_3 \neq 0$ ismeretében (4) x_{03} -at egyértelműen megadja:

$$x_{03} = \frac{P_{03}}{\Delta m_3 \cdot \omega^2}.$$

Az x_{03} kifejezésében szereplő $\frac{1}{\Delta m_3 \cdot \omega^2}$ szorzónak rugóállandómértékben megfelelő távolságot igen egyszerű, s az eddig végzett szerkesztésbe jól beleillő szerkesztéssel is meghatározhatjuk (szaggatott vonalak a 2a ábrán).

Ha \vec{BJ} a berajzolt pozitív iránynyal ellenkezőnek adódik, Δm_3 is negatív. Az elmondottak értelemszerűen alkalmazhatók akkor is, ha a tekintetbe vett gerjesztő erő valamelyik szélső tömeget támadja.

A szerkesztés — egyszerűsége folytán — igen előnyösen alkalmazható különböző típusú rezonanciagörbék (frekvencia-munkadiagramok) pontról-pontra történő megszerkesztésére is.

Az ismertett szerkesztő eljárásnak megfelelő számítás alkalmazásával — az egyszerűség és gyorsaság megmaradása mellett — a pontosság tetszés szerint fokozható.

² Ez az összefüggés a kinetikai szemléletből is kiadódik. Úgy tekinthetjük ugyanis, hogy az eredeti rendszert támadó gerjesztő erő m_3 -ra éppen akkora erőhatást fejt ki, mint amekkorát a hozzákapcsoltnak gondolt Δm_3 fejtene ki rá, azaz a gerjesztő erő Δm_3 -at kinetikai szempontból mindenben helyettesíti. Newton 2. axiómája alapján, ha a szokásnak megfelelően x_3 -ra és a gerjesztő erőre egymással megegyező pozitív irányt veszünk fel:

$$-\Delta m_3 \cdot \ddot{x}_3 = P_{03} \cdot \sin \omega t, \text{ vagyis } \Delta m_3 \cdot \omega^2 \cdot x_{03} = P_{03},$$

s ebből csakugyan

$$x_{03} = \frac{P_{03}}{\Delta m_3 \omega^2}$$

adódik.

3. Az eljárás alkalmazhatóságának kiterjesztése általános lineáris lengőrendszerre

Az általános lineáris lengőrendszer gerjesztett lengésének mátrix-differenciálegyenlete:

$$(5) \quad \mathbf{E}^0 \cdot \ddot{\mathbf{x}} = -\mathbf{U}^0 \cdot \mathbf{x} + \mathbf{p}^0,$$

itt \mathbf{E}^0 és \mathbf{U}^0 általános szimmetrikus mátrixok, főátlójukban pozitív elemekkel, s \mathbf{E}^0 pozitív definit. Erre a rendszerre a 2. pontban leírt számítás és szerkesztés közvetlenül nem alkalmazható. Ahhoz, hogy a szerkesztést ilyen lengőrendszerekre is alkalmazhassuk, a lengőrendszer mozgását leíró differenciálegyenletrendszert egy vagy több (1) alakú differenciálegyenletrendszerre kell transzformálni. A transzformációt például az [5] által adott véges iterációs eljárással végezhetjük.

4. A gerjesztett lengések elmélete néhány tételének szemléletes bemutatása

Ismeretes, hogy olyan 1. ábra szerinti lengőrendszerrel, amelynek csak a j -edik tömegét támadja ω körfrekvenciájú szinuszos gerjesztő erő, a j -edik, $(j+1)$ -edik és így tovább ... n -edik tömeg nyugalomban maradhat állandósult állapotban, ha a j -edik tömegnél befogottnak képzelte, s az első $(j-1)$ tömegből álló lengőrendszernek ω az egyik sajátkörfrekvenciája. Ez az ún. *részleges lengés* jelensége; gyakorlatilag pl. veszteségmentes lengésszigetelésre hasznosítható (dinamikus lengéscsökkentő). A 3. ábra az ismertetett szerkesztés segítségével ötszabadságfokú, a 3. tömegén gerjesztett lengőrendszerre közvetlenül szemlélteti ezt a tételt; könnyen belátható, hogy a szerkesztésnek ebben a tekintetben lényeges eredményét sem a szabadságfok, sem az egyes rugóállandók, illetve tömegek nagyságai, sem j értéke nem befolyásolják.

Azt a feltételt, hogy a j -edik tömegnél befogottnak képzelte, s az első $(j-1)$ tömegből álló lengőrendszernek ω az egyik sajátfrekvenciája, azzal elégítettük ki, hogy olyan m_1 , c_{12} , m_2 és c_{23} mérőszámokat vettünk fel, hogy a Kohn-szerkesztés baloldalt elindított törtvonala c_{23} képének jobboldali végpontjába fusson be. (Egyszerűség kedvéért úgy szerkesztettük ezt a részlengőrendszert, hogy ω a *legnagyobb* sajátkörfrekvenciája legyen.) A 2a ábra szerinti szerkesztés értelemeszerű alkalmazásával J a végesben levőnek, B végtelen távolinak, s így Δm_3 is végtelennek adódik. Ennek az a következménye, hogy (4) szerint $-P_{03} \neq 0$ lévén $x_{03}=0$ lesz, azaz a 3. tömeg mozdulatlan marad. Ennek következtében x_{03} -al nem lehet osztani, s így nem vihetjük át a 2. pontban vázolt módon a 3. tömeg mozgásegyenletének jobboldalát a baloldalra. Ez azonban eredményeink helyességét nem befolyásolja.

A 2b ábra szerkesztésének alkalmazása $x_{03}=0$ miatt $x_{04}=x_{05}=0$ eredményt szolgáltat.

x_{01} és x_{02} egyelőre határozatlanok. x_{02} abból a feltételből határozható meg, hogy a nyugalomban levő, és jobboldalról is erőmentes rugóhoz kapcsolódó 3. tömeg csak akkor lehet nyugalomban, ha a baloldalt hozzá-

kapcsolódó rugó által rá gyakorolt erő minden pillanatban kiegyensúlyozza a $P_{03} \cdot \sin \omega t$ gerjesztést, azaz ha

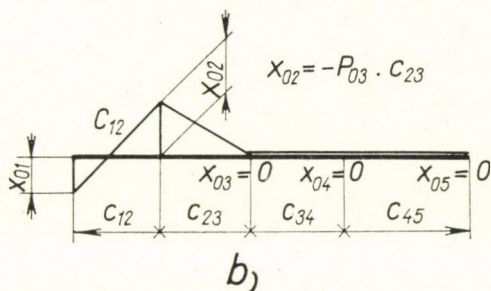
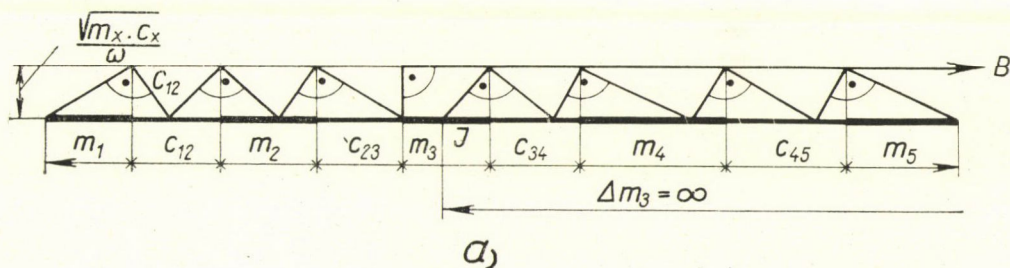
$$x_{02} = -P_{03} \cdot c_{23}.$$

Ezután x_{01} a 2b ábra szerinti módon adódik.

Figyelemreméltó jelenségként tartja számon az irodalom az úgynevezett *látszatrezonanciát* [6]. Ez abban áll, hogy adott feltételek kielégülése esetén a legalább háromszabadságfokú lengőrendszer valamennyi tömege véges amplitúdóval leng, annak ellenére, hogy valamelyik sajátkörfrekvenciájával megegyező körfrekvenciájú gerjesztő erők támadják. Csupán azt az esetet vizsgáljuk, amelyben csak egyetlen tömeget támad a rendszer valamelyik sajátkörfrekvenciájával megegyező körfrekvenciájú gerjesztő erő. Ilyen esetek szuperpozíciójával arra az általánosabb esetre is következtethetünk, amelyben valamennyi tömeget egymással megegyező körfrekvenciájú gerjesztő erők támadják.

A 2. és 3. ábrával kapcsolatban nyert eredmények alapján belátható, hogy egyetlen tömeget támadó, és e rendszer egyik a sajátkörfrekvenciájával megegyező körfrekvenciájú gerjesztő erő esetén csak akkor nem adódnak végtelen amplitúdók, ha az adott sajátkörfrekvenciához tartozó Kohn-szerkesztés a gerjesztett tömeg képét a végtelenben osztja két részre. A 4a ábra egy ilyen, négyszabadságfokú, 2. tömegben gerjesztett lengőrendszer képét és a megfelelő, balról elindított Kohn-törtvonalat mutatja. A B pont most a jobboldali, a J a baloldali végtelen távoli pontban van, így Δm_2 végtelen.

A 4b ábra lengésképén $x_{02} = 0$, azaz a 2. tömeg nyugalomban marad állandósult állapotban. x_{01} és x_{03} meghatározására csak a 2. tömeg egyensúlyát kifejező



3. ábra

$$P_{02} + \frac{1}{c_{12}} x_{01} + \frac{1}{c_{23}} x_{03} = 0$$

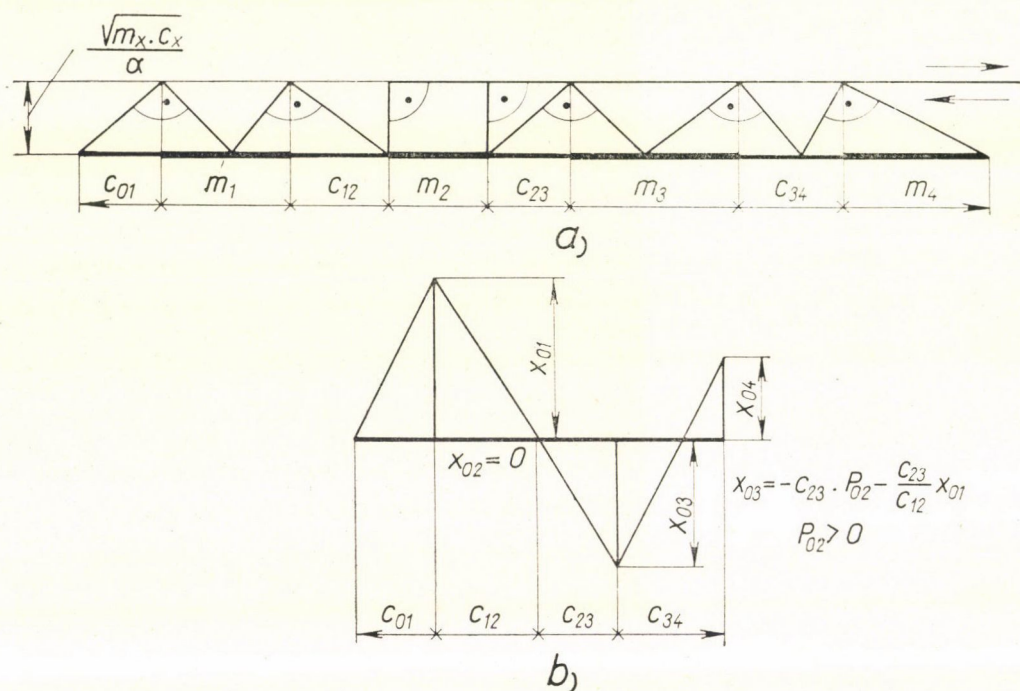
egyetlen egyenlet áll rendelkezésre. x_{01} és x_{03} meghatározása csak akkor lehetséges, ha a fizikai körülmények valami további feltétel kielégülését is megkívánják. Ha például x_{01} -et megszabjuk, $x_{03} = -c_{23} P_{02} - \frac{c_{23}}{c_{12}} x_{01}$ adódik.

A 4b ábrát ennek, és $P_{02} > 0$ -nak megfelelően rajzoltuk meg. x_{01} és x_{03} meghatározása után a többi (példánkban már csak egy ilyen van) amplitudó az ismert szerkesztéssel adódik.

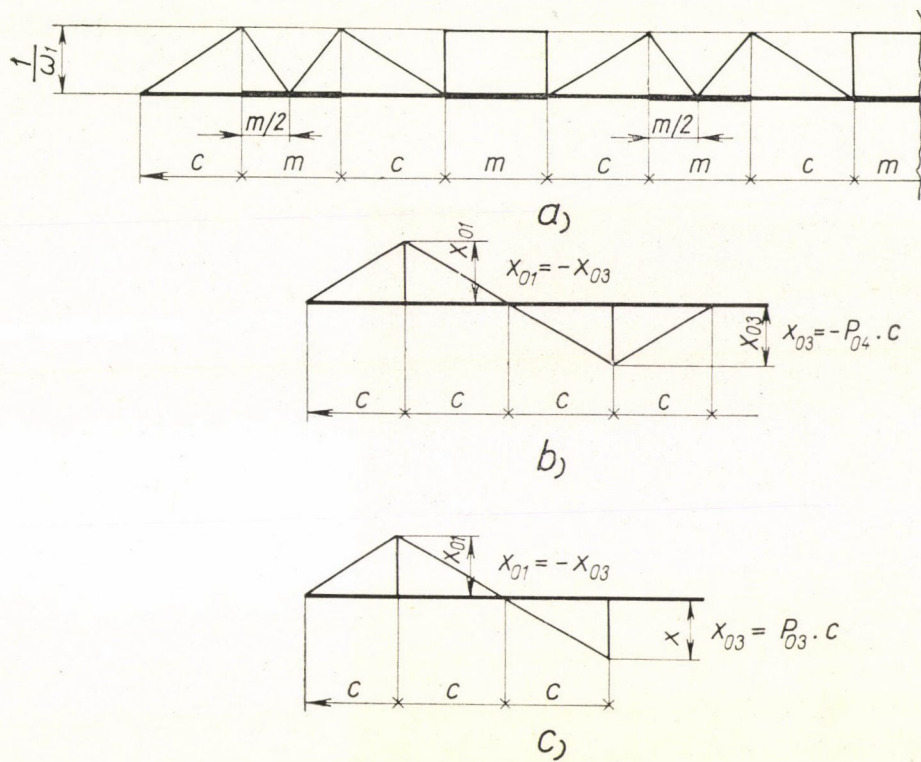
Eredményesen alkalmazható az ismertetett szerkesztés abban a különleges esetben is, ha a lengőrendszer homogén, azaz ha $m_i = m$ és $c_{ij} = c$. A következőkben néhány ilyen vizsgálatot mutatunk be. Az itt következő vizsgálat eredményeit egy általánosabb esetre érvényes jelenség felismerése és vizsgálata kapcsán Rózsa [8] a mátrixszámítás módszeres alkalmazásával nyerte.

Az 5. ábra egyik végén kötött, szabad végén $P_0 \cdot \sin \omega_1 t = P_0 \cdot \sin \frac{1}{\sqrt{\frac{c}{m}}} t$

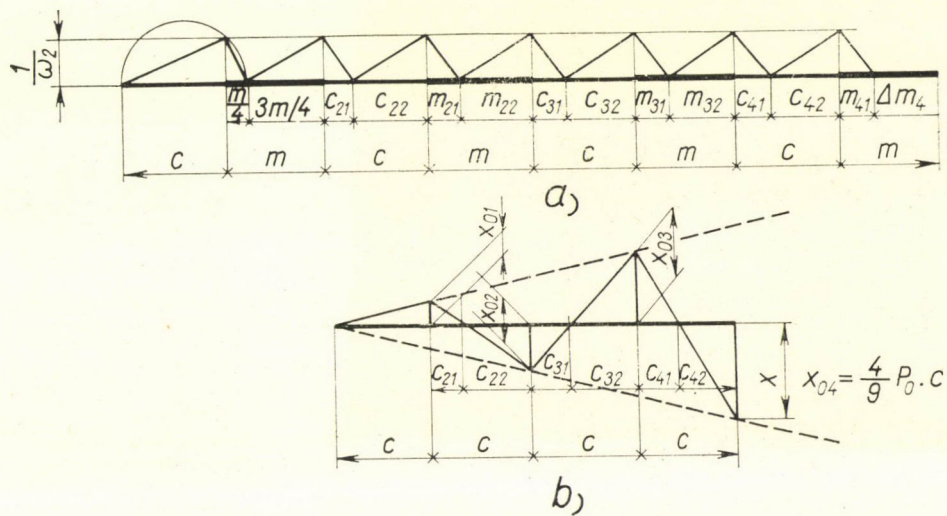
erővel gerjesztett rendszer amplitudóeloszlásának szerkesztését szemlélteti. A vázra rajzolt törtvonal (5a ábra) világosan mutatja, hogy minden második tömeg csomópontba esik, azaz nyugalomban marad. Ennek a következménye az is, hogy az amplitudók abszolút értékei megegyeznek egymással. Ha tehát



4. ábra



5. ábra



6. ábra

páros számú tömegből áll a rendszer, a végső tömeg a gerjesztés ellenére nyugalomban marad. Az 5b ábra példaképpen négyszabadságfokú rendszeren részletezi a szerkesztést. x_{03} -at a 3b ábrával kapcsolatban mondottak szerint határoztuk meg. Páratlan számú tömeg esetét háromszabadságfokú rendszeren az 5c ábra magyarázza. Itt a 2. pontban szerepelt $\Delta m_3 = m/2$ és így (4) alapján

$$x_{03} = \frac{P_{03}}{\Delta m_3 \cdot \omega_1^2} = P_{03} \cdot c.$$

Az előző feladatban ω_1 helyett $\omega_2 = \frac{1}{\sqrt{c \frac{m}{4}}}$ körfrekvenciát véve a 6.

ábrán látható eredmény adódik. Itt az eredmény szempontjából közömbös, hogy a szabadságfok páros vagy páratlan. Az ábrán példaképpen négyszabadságfokú rendszert vizsgálunk. A 6a ábrából következik, hogy általában:

$$c_{k1} = \frac{k-1}{2k-1} c$$

és

$$m_{k2} = \frac{2k+1}{4k} m.$$

Példánkban (4) alapján

$$x_{04} = \frac{P_{04}}{\Delta m_4 \cdot \omega_2^2} = \frac{4}{9} P_{04} c,$$

s általában:

$$x_{0k} = \frac{P_{0k}}{m_{k2} \cdot \omega_2^2} = \frac{k}{2k+1} P_{0k} c.$$

A 6b ábrából pedig ugyancsak elemi módon kapjuk, hogy általában

$$|x_{0k}| = \frac{c_{k2}}{c_{k1}} |x_{0,k-1}| = \frac{k}{k-1} |x_{0,k-1}|.$$

Ennek következtében

$$|x_{0k}| - |x_{0,k-1}| = 1,$$

azaz az amplitudók a gerjesztés helyétől számítva lineárisan csökkennek, tehát a 6b ábra szaggatottan rajzolt vonalai egyenesek.

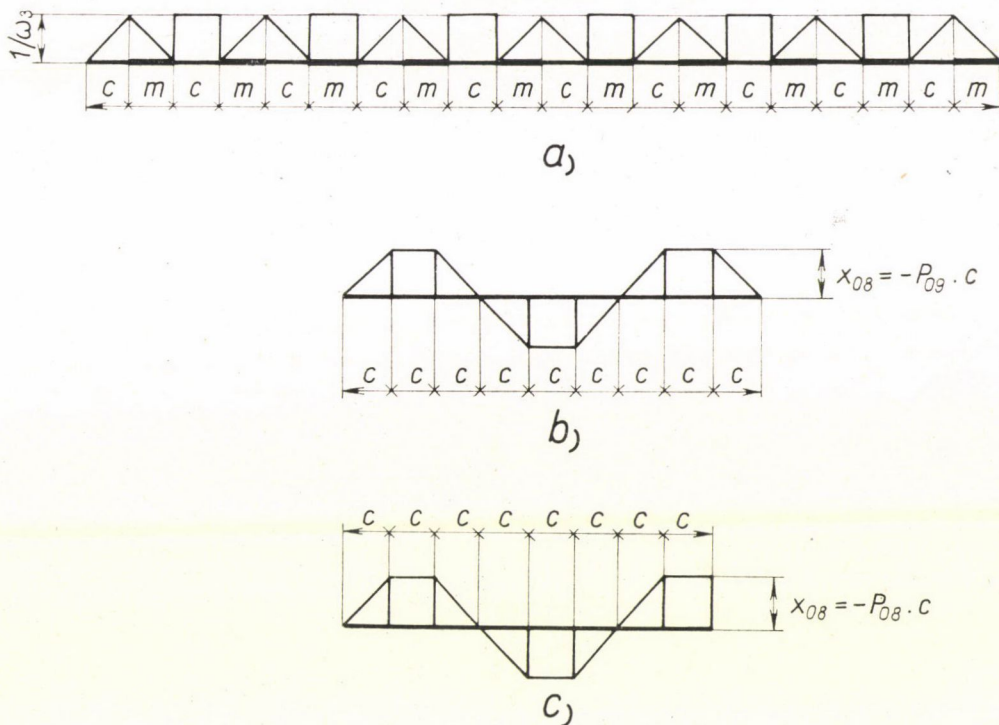
Az n indexű tömegre ($k = n$) minden határon túl növekvő szabadságfok (vagyis $n \rightarrow \infty$) esetére a fenti képletekből a következők adódnak:

$$\lim_{n \rightarrow \infty} m_{n2} \equiv \lim_{n \rightarrow \infty} \Delta m_n = \frac{1}{2} m,$$

$$\lim_{n \rightarrow \infty} x_{0n} = \frac{1}{2} P_{0\infty} \cdot c,$$

$$\lim_{n \rightarrow \infty} \frac{|x_{0n}|}{|x_{0,n-1}|} = 1.$$

Vizsgáljuk meg még azt az esetet, amelyben a gerjesztő erő körfrekvenciája $\omega_3 = \frac{1}{\sqrt{cm}}$. A szerkesztés törtvonala ekkor a 7a ábrán látható módon alakul. Ebből kitűnik, hogy $3l - 2$ (itt $l = 1, 2, 3, \dots$) szabadságfokú modellnél $\Delta m = 0$, azaz rezonancia-jelenséget kapunk. Ha a szabadságfok $3l$, $\Delta m = \infty$, azaz a gerjesztő erővel terhelt tömeg nyugalomban marad, ha pedig a szabadságfok $3l - 1$, akkor $\Delta m = -m$. Az eddigiek alapján könnyen megérthető eredményeket a 7b és 7c ábrák tartalmaznak; az előbbi egy $3l$, az utóbbi egy $3l - 1$ szabadságfokú modellen.

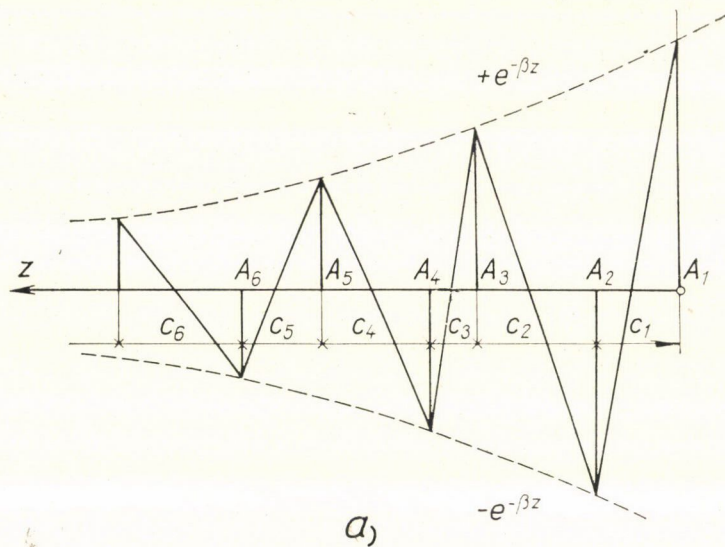


7. ábra

5. Előírt gerjesztett lengésképpel bíró lengőrendszer szerkesztése

A fentiekben ismertetett eljárás segítségével *előírt* gerjesztett lengés amplitudó eloszlással bíró lengőrendszert is szerkeszthetünk.

A 8. ábra a kötött vég felé exponenciális függvény szerint csökkenő amplitudóeloszlással bíró, egyik végén kötött, másik végén szabad és itt gerjesztett lengőrendszer paramétereinek szerkesztését mutatja be. Feladatunk szoros rokonságban van a ROUTH által vizsgált gerjesztett lengés problémával, az ún. Routh-féle jelenséggel [7]. A feladat nem egyértelműen határozott, így további kikötéseket tehetünk. Írjuk elő például, hogy a szóban forgó gerjesztett lengésképpben minden csomópont valós legyen, azaz hogy a szomszédos tömegek amplitúdói ellenkező előjelűek legyenek. Állapodjunk meg



továbbá abban, hogy az exponenciális függvény független változója a szabad végen levő, pontszerűnek veendő tömegtől számított távolság legyen, s tekintsük ezt arányosnak a rugóállandókban mért távolsággal.

A 8a ábrából kitűnik, hogy a megadott feltételeket a rugók képeinek végpontjait jelentő A_1, A_2, \dots, A_n pontok tetszés szerinti felvételével kielégíthetjük. A 8b ábra azt szemlélteti, hogy a baloldali c_{01} rugóállandót kivéve egymással megegyező rugóállandókat is felvehetünk; az exponenciális függvény tulajdonságából következik, hogy ebben az esetben a csomópont mindegyik rugóképet úgy metszi ketté, hogy a „baloldali” részek egymással egyenlők és a „jobboldali” részek is egymással egyenlők. A továbbiakban ezzel az utóbbi felvétellel dolgozunk tovább. A 8c ábra szemlélteti, hogy a jobb szélső tömeget kivéve valamennyi tömeget is egymással megegyezőnek vehetjük; ennél a választásnál a legcélszerűbb, ha a tömegnek megfelelő távolságdarabokat a c rugóállandónak megfelelő távolságdarabbal megegyező hosszúságúnak vesszük. A 8c ábra ezt a lehetőséget és az n -edik tömeg amplitudójának számításához szükséges Δm_n -et mutatja. Elemi meggondolással kiadódik, hogy ebben a legutóbbi esetben a tömegek osztópontja ugyanúgy kell, hogy elhelyezkedjék, mint a rugókon a csomópontok. Ebből a 8c ábrán látható módon kiadódik az előírt csomópontokat létesítő gerjesztő körfrekvencia reciprok értékével arányos „magasság” is. Utolsó lépésként az m_x tömeg-, a c_x rugóállandó- és az ω_x [cm/sec⁻¹] körfrekvencia-léptékek közül kettőt tetszés szerint felvehetünk, a harmadik a felvettekből az $1/\omega_x = \sqrt{m_x \cdot c_x}$ kapcsolatból számítható, és ezzel a kitűzött feladat egy lehetséges megoldása rendelkezésünkre áll.

(Beérkezett: 1959. május 15.)

IRODALOM

- [1] RAUSCH, E.: *Ingenieur Archiv* **2** (1930) 203.
- [2] KOHN, P.: „Zeichnerisches Verfahren zur Bestimmung der Torsionsschwingungen von Wellen”. *Maschinenbau* **5** (1926) 220.
- [3] BOSZNAY Á.: „Többszabadságfokú lengőrendszerek sajátkörfrekvencia- és gerjesztett lengéskép meghatározásának egy módszeréről, és annak gyakorlati alkalmazásáról”. Kandidátusi disszertáció. Budapest (1956).
- [4] SÁLYI I.: „Csillapított rendszerek gerjesztett lengései”. *A Nehézipari Műszaki Egyetem Közleményei*, Miskolc **1** (1957) 327.
- [5] FALK, S.: „Die Abbildung eines allgemeinen Schwingungssystems auf eine einfache Schwingerkette”. *Ingenieur Archiv* **23** (1955) 314.
- [6] HOLZER, H.: „Die Beseitigung der Resonanzgefahr”. *Schweizerische Bauzeitung* **82** (1923) 310.
- [7] ROUTH, E. J.: *A treatise on the dynamics of a system of rigid bodies*. Macmillan, London (1884) 230–232.
- [8] RÓZSA P.: „Elastikusán kapcsolt korpuszikuláris rendszerek kis rezgéseinek vizsgálata a mátrixszámítás alkalmazásával”. *A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei* **2** (1953) 51.
- [9] HOLZER, H.: „Torsionsschwingungen von Wellen mit beliebigen Massen”. *Schiffbau* **8** (1907) 823.

EINE METHODE ZUR UNTERSUCHUNG DER ERZWUNGENEN SCHWINGUNGEN VON SCHWINGUNGSSYSTEMEN MIT MEHREREN FREIHEITSGRADEN UND DEREN ANWENDUNGEN

von

Á. BOSZNAY

Zusammenfassung

Die Arbeit gibt eine — auf der (ursprünglich für die Bestimmung von Eigenfrequenzen von Schwingungssystemen mit mehreren Freiheitsgraden dienenden) Kohnschen Konstruktion [2] beruhende — Konstruktion zur Bestimmung der erzwungenen Schwingungsausschläge von Schwingungssystemen. Die Konstruktion kann auch in extremen Fällen als z. B. beim Phänomen der sogenannten partialen Schwingung und der Scheinresonanz angewendet werden.

Den Gedanken, dass eine Methode für die Bestimmung von Eigenfrequenzen auch für die Bestimmung von erzwungenen Schwingungsformen angewendet werden kann, findet man schon bei HOLZER [9]; eine solche Anwendung der Kohnschen Konstruktion — nach dem Wissen des Verfassers — wurde erstmals vom Verfasser [3] aufgeworfen.

Die Arbeit demonstriert ferner, dass die vorgeschlagene Methode auch zur Lösung des umgekehrten Problems (die Konstruktion eines Schwingungssystems mit vorgeschriebener erzwungener Schwingungsform) anwendbar ist.

ОДИН МЕТОД ИССЛЕДОВАНИЯ ВОЗБУЖДЕННЫХ КОЛЕБАНИЙ КОЛЕБЛЮЩИХСЯ СИСТЕМ С НЕСКОЛЬКИМИ СТЕПЕНЯМИ СВОБОДЫ И НЕСКОЛЬКО ЕГО ПРИЛОЖЕНИЙ

Á. BOSZNAY

Резюме

Работа даёт построение для определения форм возбужденных колебаний колеблющихся систем с несколькими степенями свободы, основанное на построении Конн-а [2] для определения собственных круговых частот колеблющихся систем с несколькими степенями свободы. Построение может наглядно указывать на некоторые явления, встречающиеся в особых случаях, например, явление частичного колебания и кажущегося резонанса.

Идея применения метода определения собственных круговых частот и для определения форм возбужденных колебаний встречается уже у HOLZER-а [9]; насколько автору известно, такое применение метода было впервые предложено им.

В работе, далее, показывается, что предложенный метод может применяться и к решению обратной проблемы (построение системы колебания с заданной формой возбужденных колебаний).

A WILCOXON-STATISZTIKA KÉT MÓDOSÍTÁSÁRÓL

CSÁKI ENDRE

Bevezetés

Két minta összehasonlítására jól ismert próba a Wilcoxon-próba, mely a következő statisztikán alapul: legyen $\xi_1, \xi_2, \dots, \xi_n$, ill. $\eta_1, \eta_2, \dots, \eta_m$ két egymástól független statisztikai sokaságból vett minta. Tekintsük a (ξ_i, η_j) párok közül azokat, melyekre $\eta_j < \xi_i$ ($i = 1, 2, \dots, n$; $j = 1, 2, \dots, m$). Ezek száma legyen U . Jelöljük a ξ változók közös eloszlásfüggvényét $F(x)$ -el, az η változók közös eloszlásfüggvényét $G(x)$ -el. Legyenek ezek folytonosak. A $H_0: F(x) \equiv G(x)$ hipotézist visszautasítjuk, ha U túl kicsi, vagy túl nagy, ellenkező esetben elfogadjuk.

A Wilcoxon-próba azonban, mint már MANN és WHITNEY [4] valamint D. VAN DANTZIG [5] kimutatta, csak olyan alternatív ellenhipotézissel szemben konzisztens, melyre $p = \int G(x) dF(x) \neq \frac{1}{2}$, vagyis a próba nem mindig azt dönti el, hogy $F(x)$ különbözik-e $G(x)$ -től, vagy nem, hanem csak azt, hogy $p = \frac{1}{2}$, vagy nem. Ezért természetes törekvés úgy módosítani a Wilcoxon-próbát, hogy az minden $F(x) \neq G(x)$ hipotézissel szemben konzisztens legyen. Ilyen módosításokat LEHMANN [6] és RÉNYI [7] adott meg. Az alábbi 1. §-ban kimutatjuk, hogy a két módosítás lényegében azonos. Ez egy LEHMANN—SCHEFFÉ lemma közvetlen következménye, de közvetlenül is belátható. A 2. §-ban meghatározzuk a statisztika szórását, melyre vonatkozólag a következő egyszerű formulához jutunk:

$$D^2(V) = \frac{(n+m+1)(n+m-2)}{45 \binom{n}{2} \binom{m}{2}},$$

ahol V a LEHMANN-statisztika.

1. §. A V és W statisztika

LEHMANN a következő módosítást javasolja [6]: a (ξ_i, η_j) párok helyett tekintsük a $(\xi_i, \xi_k; \eta_j, \eta_l)$ számnégyeseket, ahol $i \neq k$, $j \neq l$. Legyen V_1 azon számnégyesek száma, melyekre $\max(\xi_i, \xi_k) < \min(\eta_j, \eta_l)$, V_2 pedig azoké, melyekre $\min(\xi_i, \xi_k) > \max(\eta_j, \eta_l)$. A továbbiakban a $\max(x_1, x_2) <$

$< \min(y_1, y_2)$ eseményt $\{x_1, x_2 < y_1, y_2\}$ -vel jelöljük. A statisztika a következő:

$$(1) \quad V = \frac{V_1 + V_2}{\binom{n}{2} \binom{m}{2}}.$$

RÉNYI módosítása a következő [7]: Tekintsük a (ξ_i, η_j, η_l) számhármassok közül (ahol $j \neq l$) azokat, melyekre $\eta_j < \xi_i, \eta_l < \xi_i$. Legyen ezeknek a száma W_1 . Tekintsük a (ξ_i, ξ_k, η_j) számhármassok közül ($i \neq k$) azokat, melyekre $\xi_i < \eta_j, \xi_k < \eta_j$. Ezek száma legyen W_2 . A W statisztika a következő:

$$(2) \quad W = \frac{W_1}{n \binom{m}{2}} + \frac{W_2}{m \binom{n}{2}}.$$

A V és W statisztika között a következő egyszerű lineáris kapcsolat van:

$$(3) \quad V = 2W - 1.$$

Számítsuk ki először a V , ill. W valószínűségi változó várható értékét.

$$\begin{aligned} \mathbf{M}(V) &= \mathbf{P}(\{\xi, \xi' < \eta, \eta'\} + \{\eta, \eta' < \xi, \xi'\}) = \\ &= \int (1-F)^2 dG^2 + \int (1-G)^2 dF^2 = 2 + \int (F^2 dG^2 + G^2 dF^2) - 4 \int FG d(F+G) = \\ &= 2 + \int d(F^2 G^2) - 4 \int FG d(F+G) = \\ &= 3 - 2 \int [(F+G)^2 - (F-G)^2] d\left(\frac{F+G}{2}\right) = \frac{1}{3} + 2 \int (F-G)^2 d\left(\frac{F+G}{2}\right). \end{aligned}$$

Míg

$$\begin{aligned} \mathbf{M}(W) &= \mathbf{P}(\eta, \eta' < \xi) + \mathbf{P}(\xi, \xi' < \eta) = \int F^2 dG + \int G^2 dF = \\ &= \int (F+G)^2 d(F+G) - \int F^2 dF - \int G^2 dG - 2 \int FG d(F+G) = \\ &= \frac{8}{3} - \frac{1}{3} - \frac{1}{3} - \int [(F+G)^2 - (F-G)^2] d\left(\frac{F+G}{2}\right) = \frac{2}{3} + \int (F-G)^2 d\left(\frac{F+G}{2}\right). \end{aligned}$$

Látható, hogy a várható értékekre fennáll az

$$\mathbf{M}(V) = 2 \mathbf{M}(W) - 1$$

összefüggés. Megjegyezzük, hogy a (3) összefüggés már ebből is következik, ha felhasználjuk a következő, LEHMANN—SCHEFFÉ lemmát (lásd [6]).

Legyen $f(F, G)$ valós funkcionál értelmezve minden folytonos $F(x)$, $G(x)$ eloszlásfüggvényre. Akkor legfeljebb egy olyan $t_{n,m}$ függvény létezik, hogy $t_{n,m}(\xi_1, \xi_2, \dots, \xi_n; \eta_1, \eta_2, \dots, \eta_m)$ szimmetrikus az első n és utolsó m változójára és torzítatlan becslése $f(F, G)$ -nek minden folytonos $F(x)$, $G(x)$ -re. Ha egy ilyen $t_{n,m}$ függvény létezik és véges szórása van, akkor ennek $f(F, G)$ minden torzítatlan becslése között legkisebb a szórása.

E lemma első része azt mondja ki, hogy egy $f(F, G)$ -nek lényegében csak egy szimmetrikus torzítatlan becslése van. Így $\mathbf{M}(V)$ -nek és $\mathbf{M}(2W - 1)$ -nek szimmetrikus torzítatlan becslései azonosak, azaz fennáll a (3) összefüggés.

A (3) relációt azonban közvetlenül is be lehet látni a LEHMANN—SCHEFFÉ lemma felhasználása nélkül. Azt fogjuk bizonyítani, hogy

$$(4) \quad V_1 + V_2 = (n-1) W_1 + (m-1) W_2 - \binom{n}{2} \binom{m}{2}.$$

Legyen

$$\beta(x_1, x_2; y_1, y_2) = \begin{cases} 1, & \text{ha } \{x_1, x_2 < y_1, y_2\}, \text{ vagy } \{y_1, y_2 < x_1, x_2\} \\ 0 & \text{különben,} \end{cases}$$

továbbá

$$\gamma(x; y_1, y_2) = \begin{cases} 1, & \text{ha } y_1 < x \text{ és } y_2 < x, \\ 0 & \text{különben.} \end{cases}$$

Ekkor nyilvánvalóan

$$V_1 + V_2 = \sum_{(i,k)} \sum_{(j,l)} \beta(\xi_i, \xi_k; \eta_j, \eta_l),$$

valamint

$$W_1 = \sum_{(j,l)} \sum_{i=1}^n \gamma(\xi_i; \eta_j, \eta_l),$$

$$W_2 = \sum_{(i,k)} \sum_{j=1}^m \gamma(\eta_j; \xi_i, \xi_k),$$

ahol $\sum_{(i,k)}$, ill. $\sum_{(j,l)}$ azt jelenti, hogy az összegezés kiterjesztendő az $1, 2, \dots, n$, ill. az $1, 2, \dots, m$ elemekből alkotott összes (i, k) , ill. (j, l) elempárra, ahol $i < k$, ill. $j < l$.

Így

$$(n-1) W_1 + (m-1) W_2 = \sum_{(j,l)} (n-1) \sum_{i=1}^n \gamma(\xi_i; \eta_j, \eta_l) + \sum_{(i,k)} (m-1) \sum_{j=1}^m \gamma(\eta_j; \xi_i, \xi_k).$$

Ez a nyilvánvaló

$$(n-1) \sum_{i=1}^n a_i = \sum_{(i,k)} (a_i + a_k)$$

azonosság felhasználásával a következőképpen alakítható át:

$$\begin{aligned} & \sum_{(j,l)} (n-1) \sum_{i=1}^n \gamma(\xi_i; \eta_j, \eta_l) + \sum_{(i,k)} (m-1) \sum_{j=1}^m \gamma(\eta_j; \xi_i, \xi_k) = \\ &= \sum_{(j,l)} \sum_{(i,k)} [\gamma(\xi_i; \eta_j, \eta_l) + \gamma(\xi_k; \eta_j, \eta_l)] + \sum_{(i,k)} \sum_{(j,l)} [\gamma(\eta_j; \xi_i, \xi_k) + \gamma(\eta_l; \xi_i, \xi_k)] = \\ &= \sum_{(i,k)} \sum_{(j,l)} [\gamma(\xi_i; \eta_j, \eta_l) + \gamma(\xi_k; \eta_j, \eta_l) + \gamma(\eta_j; \xi_i, \xi_k) + \gamma(\eta_l; \xi_i, \xi_k)]. \end{aligned}$$

Fennáll a következő azonosság:

$$-1 + \gamma(\xi_i; \eta_j, \eta_l) + \gamma(\xi_k; \eta_j, \eta_l) + \gamma(\eta_j; \xi_i, \xi_k) + \gamma(\eta_l; \xi_i, \xi_k) = \beta(\xi_i, \xi_k; \eta_j, \eta_l),$$

így

$$(n-1) W_1 + (m-1) W_2 = \sum_{(i,k)} \sum_{(j,l)} [\beta(\xi_i, \xi_k; \eta_j, \eta_l) + 1] = V_1 + V_2 + \binom{n}{2} \binom{m}{2}$$

ahonnan adódik a (4) reláció.

Azt, hogy a V -statisztikán alapuló próba minden $F \neq G$ ellenhipotézissel szemben konzisztens, már LEHMANN kimutatta [6] és lényegében abból következik, hogy $D^2(V) \rightarrow 0$, midőn $\min(n, m) \rightarrow \infty$, valamint $M(V) = \frac{1}{3}$

azaz $\int (F - G)^2 d\left(\frac{F + G}{2}\right) = 0$, akkor és csak akkor áll fenn, ha $F \equiv G$.

Megjegyezzük még, hogy a V , ill. a W statisztika a következőképpen is kifejezhető:

Legyen r_1, r_2, \dots, r_n , ill. s_1, s_2, \dots, s_m azon helyek sorszáma, ahol a $\xi_1, \xi_2, \dots, \xi_n$, $\eta_1, \eta_2, \dots, \eta_m$ mintaelemek nagyság szerint rendezett egyesített sorozatában a $\xi_1^*, \xi_2^*, \dots, \xi_n^*$, ill. az $\eta_1^*, \eta_2^*, \dots, \eta_m^*$ elemek állnak.

Ekkor

$$(5) \quad V = \frac{\sum_{i=1}^n (n-i) \binom{r_i - i}{2} + \sum_{j=1}^m (m-j) \binom{s_j - j}{2}}{\binom{n}{2} \binom{m}{2}},$$

míg

$$(6) \quad W = \frac{\sum_{i=1}^n \binom{r_i - i}{2}}{n \binom{m}{2}} + \frac{\sum_{j=1}^m \binom{s_j - j}{2}}{m \binom{n}{2}}.$$

Ebből is igazolható a (3) összefüggés, csak azt kell belátni, hogy

$$\sum_{i=1}^n (i-1) \binom{r_i - i}{2} + \sum_{j=1}^m (j-1) \binom{s_j - j}{2} = \binom{n}{2} \binom{m}{2}.$$

Ez az összefüggés pedig majdnem nyilvánvaló. Azt kell belátni ugyanis, hogy a baloldalon az összes $(\xi_i, \xi_k; \eta_j, \eta_l)$ számnégyeseknek a száma áll.

2. §. A V -statisztika szórása

Míg a Wilcoxon-statisztika eloszlása lényegében ismeretes, addig a V -statisztika eloszlásáról igen keveset tudunk. LEHMANN egy tételéből ugyan következik, hogy V aszimptotikus eloszlása normális (lásd [6]), midőn $n \rightarrow \infty$ és $m \rightarrow \infty$, úgy hogy $\frac{n}{m} = C = \text{konst.}$ függetlenül attól, hogy igaz-e a nullhipotézis, vagy nem, azonban V -nek még csak a szórása sem volt eddig ismeretes. Ebben a §-ban kiszámítjuk a V statisztika szórását, midőn $F(x) \equiv G(x)$. Megjegyezzük, hogy ezen eljárás $F(x) \neq G(x)$ esetén is alkalmazható, azonban túl bonyolult formulához jutunk.

Kiindulunk abból, hogy

$$V_1 + V_2 = \sum_{(i,k)} \sum_{(j,l)} \beta(\xi_i, \xi_k; \eta_j, \eta_l),$$

így

$$\begin{aligned}
 \mathbf{M}[(V_1 + V_2)^2] &= \mathbf{M}\left[\left(\sum_{(i,k)} \sum_{(j,l)} \beta(\xi_i, \xi_k; \eta_j, \eta_l)\right)^2\right] = \\
 &= \binom{n}{2} \binom{m}{2} \left\{ \binom{n-2}{2} \binom{m-2}{2} \mathbf{M}[\beta(\xi, \xi'; \eta, \eta') \beta(\xi'', \xi'''; \eta'', \eta''')] + \right. \\
 &+ \binom{n-2}{2} 2(m-2) \mathbf{M}[\beta(\xi, \xi'; \eta, \eta') \beta(\xi'', \xi'''; \eta, \eta'')] + \\
 &+ 2(n-2) \binom{m-2}{2} \mathbf{M}[\beta(\xi, \xi'; \eta, \eta') \beta(\xi, \xi'''; \eta'', \eta''')] + \\
 (7) \quad &+ 2(n-2) 2(m-2) \mathbf{M}[\beta(\xi, \xi'; \eta, \eta') \beta(\xi, \xi'''; \eta, \eta'')] + \\
 &+ \binom{n-2}{2} \mathbf{M}[\beta(\xi, \xi'; \eta, \eta') \beta(\xi'', \xi'''; \eta, \eta')] + \\
 &+ \binom{m-2}{2} \mathbf{M}[\beta(\xi, \xi'; \eta, \eta') \beta(\xi, \xi'; \eta'', \eta''')] + \\
 &+ 2(n-2) \mathbf{M}[\beta(\xi, \xi'; \eta, \eta') \beta(\xi, \xi'''; \eta, \eta')] + \\
 &+ 2(m-2) \mathbf{M}[\beta(\xi, \xi'; \eta, \eta') \beta(\xi, \xi'; \eta, \eta'')] + \\
 &\left. + \mathbf{M}[\beta(\xi, \xi'; \eta, \eta') \beta(\xi, \xi'; \eta, \eta')]\right\},
 \end{aligned}$$

— ahol a $\xi, \xi', \xi'', \xi''', \eta, \eta', \eta'', \eta'''$ valószínűségi változók egymástól függetlenek és azonos eloszlásúak. Egyszerűség kedvéért a továbbiakban tegyük fel, hogy a fenti változók egyenletes eloszlásúak a $(0,1)$ intervallumban.

Könnyen látható, hogy

$$\mathbf{M}[\beta(\xi, \xi'; \eta, \eta') \beta(\xi'', \xi'''; \eta'', \eta''')] = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9},$$

$$\begin{aligned}
 \mathbf{M}[\beta(\xi, \xi'; \eta, \eta') \beta(\xi'', \xi'''; \eta, \eta'')] &= \mathbf{M}[\beta(\xi, \xi'; \eta, \eta') \beta(\xi, \xi'''; \eta'', \eta''')] = \\
 &= \mathbf{P}(\xi, \xi' < \eta, \eta'; \xi'', \xi''' < \eta, \eta'') + 2 \mathbf{P}(\xi, \xi' < \eta, \eta'; \xi'', \xi''' > \eta, \eta'') + \\
 &+ \mathbf{P}(\xi, \xi' > \eta, \eta'; \xi'', \xi''' > \eta, \eta'') = \\
 &= \int_0^1 \int_0^1 [1 - \max(x, y)] (1-x)(1-y) dx dy + \\
 &+ 2 \int_0^1 \int_0^y (y-x)(1-x)y dx d[1-(1-y)^2] + \\
 &+ \int_0^1 \int_0^1 \min(x, y) xy d[1-(1-x)^2] d[1-(1-y)^2] = \frac{1}{9}.
 \end{aligned}$$

$$\mathbf{M} [\beta(\xi, \xi'; \eta, \eta') \beta(\xi, \xi'''; \eta, \eta''')] = 2 \mathbf{P}(\xi, \xi' < \eta, \eta'; \xi, \xi''' < \eta, \eta''') = \\ = 2 \int_0^1 \int_0^1 \int_0^1 \min(x, y, z) \min(x, y) \min(x, z) dx dy dz = \frac{11}{90}.$$

$$\mathbf{M} [\beta(\xi, \xi'; \eta, \eta') \beta(\xi'', \xi'''; \eta, \eta')] = \mathbf{M} [\beta(\xi, \xi'; \eta, \eta') \beta(\xi, \xi'; \eta'', \eta''')] = \\ = \mathbf{P}(\xi, \xi' < \eta, \eta'; \xi'', \xi''' < \eta, \eta') + 2 \mathbf{P}(\xi, \xi' < \eta, \eta'; \xi'', \xi''' > \eta, \eta') + \\ + \mathbf{P}(\xi, \xi' > \eta, \eta'; \xi'', \xi''' > \eta, \eta') = \int_0^1 x^4 d[1 - (1 - x)^2] + \\ + 2 \int_0^1 \int_0^y (y - x)^2 dx^2 d[1 - (1 - y)^2] + \int_0^1 x^2 d[1 - (1 - x)^4] = \frac{7}{45},$$

$$\mathbf{M} [\beta(\xi, \xi'; \eta, \eta') \beta(\xi, \xi'''; \eta, \eta')] = \mathbf{M} [\beta(\xi, \xi'; \eta, \eta') \beta(\xi, \xi'; \eta, \eta''')] = \\ = \mathbf{P}(\xi, \xi' < \eta, \eta'; \xi, \xi''' < \eta, \eta') + \mathbf{P}(\xi, \xi' > \eta, \eta'; \xi, \xi''' > \eta, \eta') = \\ = \int_0^1 x^3 d[1 - (1 - x)^2] + \int_0^1 x^2 d[1 - (1 - x)^3] = \frac{1}{5},$$

$$\mathbf{M} [\beta(\xi, \xi'; \eta, \eta') \beta(\xi, \xi'; \eta, \eta')] = \mathbf{M} [\beta(\xi, \xi'; \eta, \eta')] = \frac{1}{3}.$$

Igy tehát

$$\mathbf{D}^2(V_1 + V_2) = \mathbf{M}[(V_1 + V_2)^2] - [\mathbf{M}(V_1 + V_2)]^2 = \binom{n}{2} \binom{m}{2} \frac{(n+m+1)(n+m-2)}{45},$$

tehát

$$(8) \quad \mathbf{D}^2(V) = \frac{(n+m+1)(n+m-2)}{45 \binom{n}{2} \binom{m}{2}}.$$

Megjegyezzük még, hogy a (7) formulában szereplő várható értékek egyszerű kombinatorikus megfontolással is kiszámíthatók. A fenti módszert azonban előnyben kell részesíteni a kombinatorikussal szemben, mert ez csak abban az esetben alkalmazható, midőn $F(x) \equiv G(x)$, míg az előbbi módszerrel tetszőleges $F(x)$, $G(x)$ esetén célhoz jutunk, csak bonyolultabbak lesznek a formulák.

Határozzuk meg például kombinatorikus úton a következő várható értéket:

$$\mathbf{M} [\beta(\xi, \xi'; \eta, \eta') \beta(\xi'', \xi'''; \eta, \eta')].$$

Ez a következő (diszjunkt) események összegének valószínűsége:

$$\{\xi, \xi', \xi'', \xi''' < \eta, \eta'\}; \{\xi, \xi' < \eta, \eta' < \xi'', \xi'''\}; \{\xi'', \xi''' < \eta, \eta' < \xi, \xi'\}; \\ \{\eta, \eta' < \xi, \xi', \xi'', \xi'''\}.$$

A $\xi, \xi', \xi'', \xi''', \eta, \eta'$ változók összes lehetséges sorrendjének száma $6!$
Az első és negyedik esemény lehetőségeinek száma $4! 2!$ míg a másodiké és
harmadiké $2! 2! 2!$

Igy

$$M[\beta(\xi, \xi'; \eta, \eta') \beta(\xi'', \xi'''; \eta, \eta')] = \frac{2 \cdot 4! 2! + 2 \cdot 2! 2! 2!}{6!} = \frac{7}{45}.$$

(Beérkezett: 1959. május 28.)

IRODALOM

- [1] WILCOXON, F.: „Individual comparisons by ranking methods”. *Biometrics Bulletin* **1** (1945) 80—83.
- [2] WILCOXON, F.: „Individual comparisons of grouped data by ranking methods”. *Journ. Econ. Entomology* **39** (1946) 269.
- [3] WILCOXON, F.: „Probability tables for individual comparisons by ranking methods”. *Biometrics Bulletin* **3** (1947) 119—122.
- [4] MANN, H. B.—WHITNEY, D. R.: „On a test whether one of two random variables is stochastically larger than the other”. *The Annals of Mathematical Statistics* **18** (1947) 50—60.
- [5] VAN DANTZIG, D.: „On the consistency and the power of Wilcoxon's two sample test”. *Indagationes Mathematicae* **13** (1951) 1—8.
- [6] LEHMANN, E. L.: „Consistency and unbiasedness of certain nonparametric tests”. *The Annals of Mathematical Statistics* **22** (1951) 165—180.
- [7] RÉNYI A.: „Újabb kritériumok két minta összehasonlítására”. *A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei* **2** (1953) 243—257.

О ДВУХ ВИДОИЗМЕНЕНИЯХ СТАТИСТИКИ WILCOXON-A

E. CSÁKI

Резюме

LEHMANN [6] и RÉNYI [7] видоизменили статистику WILCOXON-a относящуюся к сравнению двух выборок, так, что она уже консистентна относительно всякой альтернативной контргипотезы. В настоящей работе доказывается, что между двумя изменёнными статистиками существует простая линейная связь. Вычисляется дисперсия статистики, что делает возможным применение пробы к большой выборке приближением с помощью нормального распределения.

ON TWO MODIFICATIONS OF THE WILCOXON-STATISTIC

E. CSÁKI

Summary

LEHMANN [6] and RÉNYI [7] have modified the Wilcoxon-statistic concerning the comparison of two samples, in a form which is consistent against every alternative hypothesis. This paper shows, that between the two modified statistics a simple linear relation exists. The dispersion of the statistic is determined, which gives in case of large sample the possibility to use this test by approximation with normal distribution.

NÉHÁNY MEGJEGYZÉS NOMOGRAMOK LEOLVASÁSI HIBÁIHOZ

HASZPRA OTTÓ¹

Nomogramok hibáinak szabatos meghatározásával csak az utolsó évtizedben kezdtek a szerzők behatóbban foglalkozni. Gondolunk itt elsősorban PENTKOVSZKIJ Nomográfiájára [1], melynek egyik központi kérdése az előírt pontosságú nomogram készítése, továbbá HAJÓS GYÖRGY és PÁL SÁNDOR cikkeire [2], [3], valamint még néhány szerzőnek az utóbbi években megjelent munkájára [4], [5], [6], [7], amelyek ugyancsak PENTKOVSZKIJ nyomdokain, ill. hasonló úton haladnak.

Ha a nomogram alakját, szerkezetét már meghatároztuk, a nomogram méretét már csak az eredményben megkívánt pontosság dönti el. Pontsoros nomogramoknál például a nomogram hibája a

$$(1) \quad \Delta z = \frac{\Delta s}{C |f'(z)|}$$

képlet alapján számítható, ahol Δz az eredmény hibája, Δs az eredmény-skálán figyelembe vett geometriai pontatlanság, $C \cdot f(z)$ pedig az eredmény-skála ívhosszát adó skálafüggvény, amelynek C nyújtási tényezője (s így az egész nomogram nagyítási mértéke is) Δs és Δz felvételétől, ill. előírásától függ. Tekintve, hogy a C nyújtási tényező, s ezzel együtt a nomogram lineáris méretei is egyenesen, a nomogram által betöltött terület pedig négyzetesen arányosak a Δs geometriai pontatlansággal, a gazdaságos helykihasználás, a kényelmesebb szerkesztés és a leolvasási hiba reális figyelembevétele érdekében nagy gondot kell fordítani Δs meghatározására.

Ellentétben az idézett szerzőkkel, akik az eredményskála geometriai pontatlanságának meghatározásában a tökéletes biztonság elvét követik, mi a valószínűségszámítási alapon álló utat tartjuk gyakorlatilag indokoltabbnak és követendőnek.

A régi módszert legtisztábban (pontsoros nomogramokra vonatkozóan és rajzi pontatlanságoktól eltekintve) PÁL SÁNDOR [2] foglalta össze, aki az eredményskála geometriai pontatlanságára a

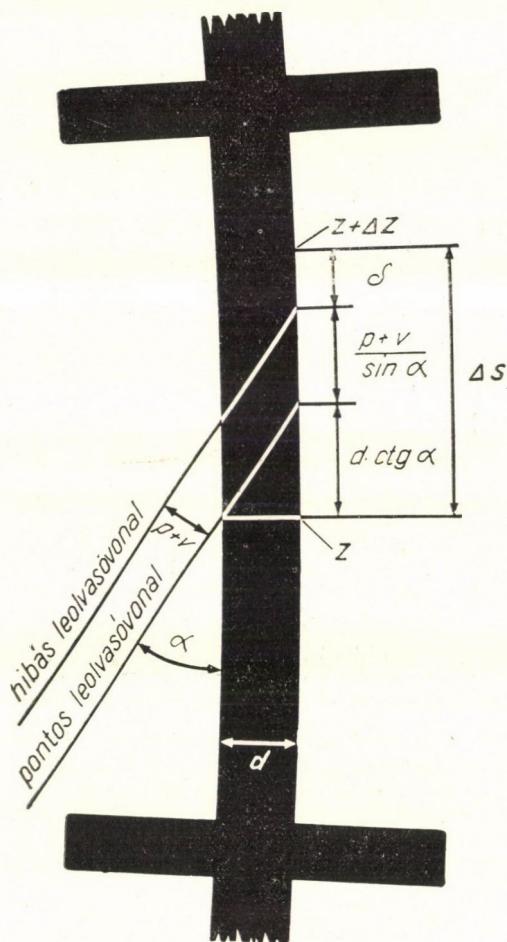
$$(2) \quad \Delta s = \frac{v + p}{\sin \alpha} + d \cdot \operatorname{ctg} \alpha + \delta$$

¹ Vizgázdálkodási Tudományos Kutató Intézet, Budapest.

képletet adta. A betűknek a jelen cikk szempontjából élesített magyarázatát az alábbiakban adjuk és az 1. ábrán figyelemmel kísérelhetjük.

Δs az eredményskála geometriai pontatlanságának felső korlátja.

v a leolvasóvonalnak az adott változók skáláin az eredményre nézve legkedvezőtlenebbül pontatlan beállítása miatt az eredményponttól való eltávolodása.



1. ábra

p a nomogram síkja és a leolvasóvonal közti véges távolság, valamint a nem merőleges ránézés miatt fellépő parallaxisnak, továbbá — szerintünk — még a leolvasóvonal véges vastagságából adódó s a leolvasóvonalra ugyancsak merőleges eltérésnek felső korlátja.

a a leolvasóvonal és az eredményskála metszési szöge (hegyesszög!).

d az eredményskála vastagsága.

δ az eredményskálán végzett interpoláció geometriai pontatlanságának felső korlátja.

A képlet két szempontból tehető kritika tárgyává.

1. Minthogy rajzi pontatlanságokat nem vesz figyelembe, a geometriai skála a valóságos és így véges d vastagságú skála közepén húzódik. A (2) egyenlet $d \cdot \operatorname{ctg} a$ tagja helyett tehát legfeljebb $\frac{d}{2} \operatorname{ctg} a$ vehető. Általában azonban

még ez sem indokolt.

2. A (2) képletben szereplő geometriai mennyiségek — a kivételével — valójában normális eloszlásúnak tekinthető, zérus várható értékű valószínűségi változók. Így korlátjuk elvileg nincs is. A felvett mennyiségek — de csak az egyenlet jobboldalán — *gyakorlati hibakorlátnak* tekinthetők, ha a gyakorlati hibakorlátot úgy definiáljuk, mint az illető mennyiség szórásának valamilyen k -szorosát (k -t szokás pl. 3-nak választani), melynek előfordulási valószínűsége közel áll az 1-hez. A baloldalon álló Δs már nem tekinthető az előbbiekkal egyenrangú gyakorlati hibakorlátnak, mert a lineáris szuperpozíció folytán előfordulásának valószínűsége közelebb áll az 1-hez, mint a jobboldalon szereplő mennyiséké. Ez azt jelenti, hogy a képlet Δs -re nagyobb értéket ad annál az értéknél, amelynek valószínűsége a felvett v , p , d és δ korlátok valószínűségeivel egyenlő volna, s amelyet az „egyenlő valószínűségű hibák elve” alapján indokoltnak tartunk meghatározni.

Véleményünk szerint tehát helyesebb az eredő geometriai pontatlanság meghatározására a

$$(3) \quad K(\Delta s) = \sqrt{\frac{K^2(v) + K^2(p)}{\sin^2 a} + K^2(t) \operatorname{ctg}^2 a + K^2(\delta)}$$

vagy a

$$(4) \quad K(\Delta s) = k \cdot D(\Delta s) = k \sqrt{\frac{D^2(v) + D^2(p)}{\sin^2 a} + D^2(t) \operatorname{ctg}^2 a + D^2(\delta)}$$

képletet használni, ahol $K(\dots)$ jelenti a zárójelben álló geometriai pontatlanságok gyakorlati hibakorlátját, $D(\dots)$ a szórását, k pedig a kettő viszonyát. A 2. ábrán v , p , t és δ tehát már nem maximális értékeket, hanem valószínűségi változók egyidejűleg előforduló értékeit jelentik, amelyek így lineárisan összegeződve adják a Δs eredő geometriai pontatlanság értékét. Δs gyakorlati hibakorlátját azonban természetesen csak a (3) vagy a (4) képlet szerinti kvadratikus szuperpozícióval indokolt meghatározni.

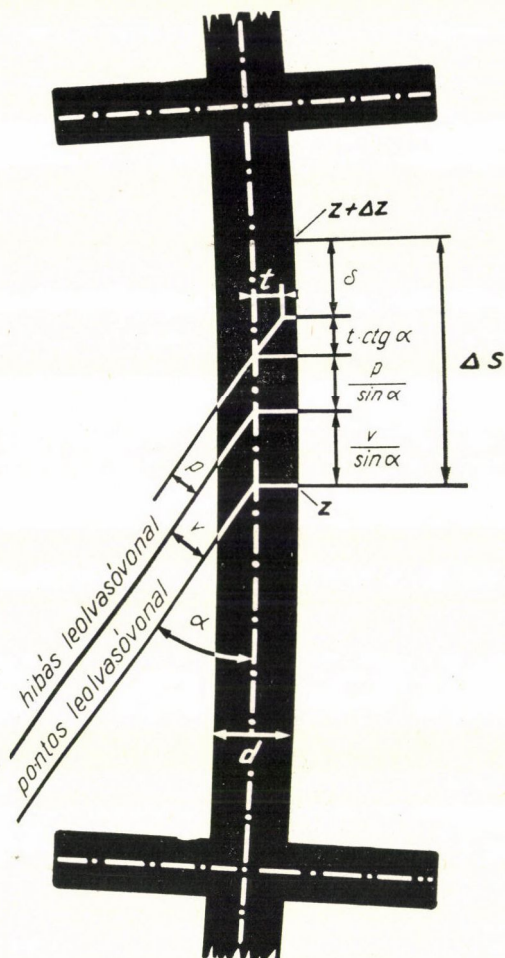
Hasonlítsuk össze a (2) és a (3) képletekből számítható eredő geometriai pontatlanságot. Az egyes geometriai pontatlanságok értékei természetesen csak tájékoztató jellegűek, hiszen nagyon függnek a skálabeosztás sűrűségétől és a nomogram használójának adottságaitól.

Legyen a (2) képletben $p \approx 0,05$ mm, $d \approx 0,3$ mm, $\delta \approx 0,1$ mm. A (3) képletben pedig ezeknek megfelelően $K(p) = p \approx 0,05$ mm, $K(t) \leq \frac{d}{2}$, tehát

$$K(t) \approx 0,1 \text{ mm}, \quad K(\delta) = \delta \approx 0,1 \text{ mm}.$$

v , ill. $K(v)$ értékét számítással kell meghatározni. Ha az eredményskála például az adott skálák között van, akkor az idézett irodalom (pl. [1], [3]) szerint v értéke egyenlő az adott skálakon elkövethető beállítási pontatlansággal. Ha az „egyenlő valószínűségű hibák elvét” alkalmazzuk, $K(v)$ az adott skálakon elkövethető beállítási pontatlanság és ennek $\frac{1}{\sqrt{2}}$ -szöröse között van.

Az előbbi azokban a pontokban éri el, ahol az eredményskála valamelyik adott skálát metszi vagy érinti, az utóbbit pedig azokban a pontokban, amelyek az adott pontok közti távolságot felezik.



2. ábra

Legyen tehát a (2) képletben $v = 0,1$ mm, akkor a (3) képletben $K(v) = 0,07 \div 0,1$ mm, amit a számításoknál vegyünk $K(v) = 0,08$ mm értékkel figyelembe. A metszési szög legyen a legkedvezőtlenebb: $\alpha = 20^\circ$.

A (2) képlet alapján

$$\Delta s = \frac{0,1 + 0,05}{\sin 20^\circ} + 0,3 \cdot \operatorname{ctg} 20^\circ + 0,1 \approx 1,36 \text{ mm}.$$

A (3) képlet alapján

$$K(\Delta s) = \sqrt{\frac{0,08^2 + 0,05^2}{\sin^2 20^\circ} + 0,1^2 \cdot \operatorname{ctg}^2 20^\circ + 0,1^2} \approx 0,4 \text{ mm}.$$

Ez tehát a maximális geometriai pontatlansággal tervezett nomogramokhoz képest azt jelenti, hogy az egyenlő valószínűségi hibák elve alapján, továbbá a skála vonalvastagságának reálisabb figyelembevételével a geometriai pontatlanságot mintegy harmadára, a nomogram által elfoglalt területet mintegy tizedére sikerül csökkenteni.

Meg kell jegyeznünk, hogy PENTKOVSKIJ [1] maga sem ad a pontsoros nomogram eredményskálájának geometriai pontatlanságára oly nagy értéket, mint amekkora a (2) képlet alapján adódott. Ezt a kisebb (0,5 mm-es) értéket azonban nem indokolja, saját tárgyalásából pedig nem következik.

Jelen tanulmány nem lép fel az újdonság közlésének igényével. Csupán régen ismert dolgok új területen való alkalmazásának előnyeire kíván rámutatni. Emiatt a háromváltozós pontsoros nomogramokon kívül nem foglalkozik a többváltozós, illetve más típusú nomogramok esetével. Ugyancsak nem foglalkozik a rajzi pontatlanságokkal és a nomogramtól független öröklött hibáknak a nomogramtól megkívánható leolvasási hibára való hatásával sem. Ezeknél az előbbi elvek alkalmazásának egyszerű gyakorlásáról van csak szó, ismertetésük inkább műszaki folyóiratba kívánczik. Nem kíván kitérni az eredő geometriai pontatlanság származásának vitatására, bár nyilvánvaló, hogy a hiba származtatásának más és talán jobb matematikai modellje is elképzelhető, mint amit v , p , t és δ alapul vételével figyelembe vettünk.

Még megjegyezzük, hogy a Descartes-féle vonalsereges nomogramokra vonatkozóan a pontatlanságok kvadratikusszerű eloszlását már alkalmazta SCHWERDT [8].

(Beérkezett : 1959. június 20.)

IRODALOM

- [1] ПЕНТКОВСКИЙ, М. В.: *Номография*. ГТТИ. Москва—Ленинград, 1949.
- [2] HAJÓS Gy.: „A nomográfia alkalmazhatóságának határaitól”. *A Magyar Tudományos Akadémia III. (Matematikai és Fizikai) Osztályának Közleményei* 1 (1951) 268—273.
- [3] PÁL S.: „Grafikus-numerikus módszer pontsoros és pontmezős nomogramok tervezésére”. *A MTA Matematikai Kutató Intézetének Közleményei* 1 (1956) 223—272.
- [4] НЕВСКИЙ, Б. А.: *Справочная книга по номографии*. ГТТИ. Москва—Ленинград, 1951.
- [5] MEYER ZUR CAPELLEN, W.: *Leitfaden der Nomographie*. Springer, Berlin—Göttingen—Heidelberg, 1953.
- [6] BAL, L.—RADÓ, F.: *Lección de nomografie*. Editura Tehnica, Bukarest, 1956.
- [7] HASZPRA O.: *Nomográfia*. Felsőoktatási Jegyzetellátó Vállalat. Budapest, 1955.
- [8] SCHWERDT, H.: *Lehrbuch der Nomographie auf abbildungsgeometrische Grundlage*. Springer, Berlin, 1924.

НЕСКОЛЬКО ЗАМЕЧАНИЙ В СВЯЗИ С ПОГРЕШНОСТЯМИ НОМОГРАММ

O. HASZPRA

Резюме

Точным определением погрешностей номограмм наиболее подробно занимались М. В. Пентковский [1] и S. PÁL [2]. Для случая створных номограмм, отвлекаясь от неточностей чертежа, они дали формулу (2) для

геометрической неточности шкалы результатов (см. чертёж 1). Фигурирующие в формуле величины они считают максимальными значениями.

В действительности эти величины — за исключением a — случайные величины с нулевым математическим ожиданием, распределение которых может считаться нормальным, следовательно, теоретически они не имеют граней. Мы можем говорить лишь о *практических гранях*, являющихся какими-нибудь k -кратными дисперсиями этих величин.

Линейная суперпозиция этих максимальных погрешностей по формуле (2) приводит к слишком большой итоговой геометрической неточности. Если мы хотим, чтобы вероятность итоговой геометрической неточности совпадала с вероятностями частичных погрешностей, фигурирующих справа, то следует применять квадратичную суперпозицию, которую дают формулы (3) и (4). Согласно чертежу 2, в этих формулах лучше принимается во внимание и влияние ширины шкалы.

В результате удалось добиться значительного сокращения размеров номограмм с заданной точностью.

EINIGE BEMERKUNGEN ZUR BERECHNUNG DER NOMOGRAMMABLESEFEHLER

O. HASZPRA

Zusammenfassung

Mit der präzisen Bestimmung der Nomogrammfehler befassten sich am ausführlichsten M. W. PENTKOWSKI [1] und S. PÁL [2], die (auf Fluchtlinientafeln bezüglich und von Ungenauigkeiten der Zeichnung abgesehen) für die geometrische Ungenauigkeit der Ergebnisskala die Formel (2) angaben (vgl. Bild 1). Sie behandeln die in der Formel enthaltenen Werte als Maximalwerte.

Eigentlich sind diese Werte — mit Ausnahme von a — als normal verteilt zu betrachtende Wahrscheinlichkeitsveränderliche mit dem Erwartungswert Null, besitzen also prinzipiell gar keine Schranken. Es kann nur von *praktischen Schranken* gesprochen werden, und zwar mit einem angenommenen k -fachen Wert der Dispersion der betreffenden Werte.

Eine der Formel (2) entsprechende lineare Superposition dieser Fehler-schranken ergibt als Resultat eine viel zu grosse geometrische Ungenauigkeit. Soll die Wahrscheinlichkeit der resultierenden geometrischen Ungenauigkeit mit der Wahrscheinlichkeit der rechts angeführten Teilfehler übereinstimmen, ist die mit Formel (3) oder (4) angegebene quadratische Superposition anzuwenden. Mit diesen Formeln wird nach Bild 2 auch die Wirkung der Skalenlinienstärke richtiger beachtet.

Solcherart gelang es eine bedeutende Abmessungsverringerung des mit der erwünschten Genauigkeit zu konstruierenden Nomogrammes zu erreichen.

RÉFRIGÉRATION DE RETOUR DES MATIÈRES GRANULEUSES¹

par

GEORGES ADLER

Introduction

On appelle réfrigération de retour la méthode de réfrigération où le médium réfrigérant et la matière à réfrigérer circulent dans des directions opposées. L'appareil le plus simple qui sert à la réfrigération de retour des liquides se compose de deux tuyaux concentriques. Le médium réfrigérant circule dans l'espace cylindrique entre les deux tuyaux, tandis que la matière à réfrigérer circule dans le tuyau intérieur dans une direction opposée au mouvement du médium réfrigérant. Les calculs se rapportant à cet appareil réfrigérant peuvent être trouvés dans de nombreux manuels (voir p. e. [3]).

Dans le cas où la matière à réfrigérer n'est pas un liquide, l'appareil réfrigérant peut être représenté par un seul tuyau, dans lequel le liquide réfrigérant et la matière à réfrigérer continuent à circuler dans des directions opposées, de telle façon que les granules de la matière à réfrigérer sont en contact immédiat avec le médium réfrigérant.

Supposons que la capacité calorifique de l'appareil servant à mouvoir la matière à réfrigérer dans le contre-courant est nulle, c'est-à-dire qu'elle est négligeable dans nos calculs. Supposons de plus que les diamètres des granules peuvent être considérés comme nuls en comparaison avec les autres données de l'appareil et que la matière à réfrigérer se répartit uniformément dans le médium réfrigérant, c'est-à-dire que la quantité de la partie de la matière à réfrigérer qui peut être trouvée dans un certain volume du médium réfrigérant est proportionnelle à la grandeur du volume en question.

Si la matière à réfrigérer avait une conductibilité thermique infinie, alors la température s'égalerait pendant une durée nulle dans les granules respectifs, et ainsi la température des granules serait une fonction de la seule coordonnée longitudinale parallèle à l'axe du tuyau formant l'appareil réfrigérant. Pour cette raison, les calculs relatifs au réfrigérant de retour servant à la réfrigération des liquides peuvent être appliqués dans ce cas aussi, sans aucune modification.

Par contre, dans le cas où la conductibilité thermique de la matière à réfrigérer est finie, nous devons tenir compte du processus de conduction de la chaleur se produisant dans les granules respectifs. Dans cet ouvrage nous voulons justement nous occuper de ce dernier cas des matières à réfrigérer de conductibilité thermique finie.

¹ Les calculs contenus dans cet ouvrage ont été effectués sur la commande de l'Institut de Recherche de l'Industrie du Froid, de la Viande et des Conserves.

Ce travail se rattache intimement au travail [1] de l'auteur. C'est pourquoi les équations resp. conditions (I), (a), (b), (c_1) ou bien (c_1^*) et (c_2) se rapportant au réfrigérant de retour sont conformes à celles de mêmes numéros du travail [1]. Il en résulte que la méthode de solution sera la même que celle de l'étude mentionnée [1]. Vu que dans ce travail [1] nous avons démontré l'existence de la solution du problème traité, et que cette démonstration peut être appliquée au problème actuel, nous ne nous occuperons pas de problèmes d'existences se rapportant au cas général, mais nous effectuerons les calculs traitant le cas le plus important du point de vue des applications. Donc le problème de la réfrigération de retour des matières granuleuses de conductibilité thermique finie peut être considéré comme un exemple se rapportant au type nouveau des problèmes aux limites traité dans le travail [1].

À l'encontre de l'analogie des problèmes traités dans le travail [1] et dans l'ouvrage présent, vu la différence de signe dans les conditions (b) (dans le travail présent β est négatif: $\beta = -1$), le principe du maximum figurant dans l'ouvrage [1] ne peut pas être appliqué dans le cas du réfrigérant de retour. Ainsi, au lieu du principe du maximum du travail [1], nous démontrerons deux autres principes du maximum correspondant à la nature du problème présent.

Dans le § 1 nous établirons les équations et les conditions caractérisant le processus de réfrigération. Dans le § 2 nous démontrerons deux principes du maximum et l'unicité de la solution du problème. Dans la partie I du § 3 nous effectuerons les calculs pour un cas concret, celui des granules de forme sphérique. Les démonstrations des assertions moins essentielles qui se présentent au cours des calculs seront effectuées dans la partie II du paragraphe, pour ne pas troubler la marche du calcul. Dans le § 4 nous ferons connaître le mode de détermination de la longueur du réfrigérant. Dans ce paragraphe, nous nous baserons en grand partie sur un point de vue purement physique en faisant abstraction aux considérations mathématiques rigoureuses. Enfin dans l'appendice nous examinerons la répartition des racines des équations (20.a) et (20.b), appelées équations caractéristiques, dans le plan des nombres complexes. Nous effectuerons cet examen séparément (d'ailleurs cet examen aurait lieu dans la partie II du § 3), parce qu'au cours de la solution des équations aux dérivées partielles à l'aide de la transformation de Laplace se présentent souvent des équations caractéristiques de type semblable à celles ci-dessus (voir p. e. [4], [7] p. 153, [9]), mais dans les manuels généralement connus on ne trouve pas la discussion des racines de ces équations. Nous sommes d'avis, que notre méthode² peut être appliqué avec succès dans d'autres cas aussi.

Les notations employées sont indiquées à la fin de l'article.

Définitions: Soit Σ un sous-ensemble de la frontière \mathfrak{B} du domaine \mathfrak{A} . Supposons la fonction $f(P)$ continue et admettant une borne supérieure sur l'ensemble $\mathfrak{A} + (\mathfrak{B} - \Sigma)$. Nous dirons alors que la fonction $f(P)$ possède une discontinuité de type **A** sur la portion Σ de la frontière. Si nous remplaçons dans cette définition la borne supérieure par la borne inférieure, nous dirons que la discontinuité est de type **A**. Dans le cas où une fonction possède simultanément une discontinuité de type **A** et de type **A** sur la portion Σ , la discontinuité sera dite de type **A**.

² L'idée de l'application du théorème de Rouché sous la forme figurant dans la démonstration est due à M. FREUD.

Nous dirons que le point P_0 de la frontière S du domaine C possède la propriété **B**, si le domaine C contient l'intérieur d'une sphère Γ , dont P_0 est sur la frontière. Dans le cas où chaque point de la frontière S possède la propriété **B**, nous dirons que la frontière S possède la propriété **B**.

On appelle surface de Liapounoff toute surface jouissant des trois propriétés suivantes (voir [5]):

1) il existe un plan tangent, et par conséquent une normale à la surface en chacun de ses points,

2) soit ϑ l'angle des normales appartenant aux points M_1 et M_2 , alors

$$\vartheta < E(\overline{M_1 M_2})^\delta \quad (0 < \delta \leq 1),$$

où E et δ sont des nombres fixés,

3) il existe un nombre $d > 0$ qui possède la propriété suivante: une droite parallèle à la normale appartenant à un point arbitraire M de la surface n'a que tout au plus un seul point commun avec la partie de la surface qui se trouve dans la sphère de centre M et de rayon d .

Nous appelons fonction calorique d'unité du domaine C la fonction $E_C^{(a)}(P, t)$ satisfaisant à l'équation de la chaleur (1) (voir [1], p. 118) pour $t > 0$ dans le domaine C limité par la surface S , de plus

1. au cas où $a > 0$, continue dans

$$(C + S) \times (t \geq 0),$$

2. au cas où $a = 0$, continue dans $[C \times (t \geq 0)] + [S \times (t > 0)]$, et ayant une discontinuité de type **A** sur la portion $S \times (t = 0)$ de la surface de ce domaine et qui satisfait à la condition initiale et à la condition aux limites suivantes:

$$E_C^{(a)}(P, 0) = 0 \quad (P \in C)$$

$$E_C^{(a)}(P, t) + a \frac{\partial E_C^{(a)}(P, t)}{\partial n} = 1 \quad (t > 0, P \in S).$$

Dans toute l'étude, nous supposons que les solutions de l'équation de la chaleur possèdent des secondes dérivées continues selon les coordonnées de lieu x_i et des premières dérivées continues selon le temps t sur l'ensemble, où elles satisfont à l'équation de la chaleur, de plus que la frontière S des granules est une surface mesurable [6].

§ 1. Établissement du problème

Le tuyau constituant l'appareil réfrigérant est installé le long de l'intervalle $(0, l)$ de l'axe x , et il est thermiquement isolé de son milieu. Les granules de la matière à réfrigérer resp. le liquide réfrigérant circulent dans la direction des valeurs croissantes resp. décroissantes de la coordonnée x . Les granules sont superposables et leur diamètre est négligeable en comparaison de la longueur et du diamètre du tuyau. Pour simplifier nous supposons les granules homogènes, restriction non essentielle en ce qui concerne la méthode de la solution. Nous nous limiterons au cas stationnaire, indépendant du temps, à savoir que la température du liquide réfrigérant resp. la répartition de la température dans les granules respectifs sont les fonctions de la seule coordonnée x . Cela signifie que les granules se trouvent continuellement dans les mêmes

circonstances thermiques extérieures et intérieures dans une certaine section transversale du tuyau.

Considérons un granule assujéti au système de coordonnées (x_1, x_2, x_3) . Désignons par C ce granule resp. le domaine qu'il occupe. Soit S la frontière du domaine C . Suivons avec attention ce granule à partir de l'instant de son entrée dans le tuyau. Mesurons le temps à partir de cet instant. Soit $u(P, t)$ la température du granule. Cette fonction u doit satisfaire à l'équation de la chaleur

$$(1) \quad \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} = a^2 \frac{\partial u}{\partial t} \quad \left(a^2 = \frac{\rho_1 \sigma_1}{k} > 0 \right)$$

dans le domaine C .

Soit $F(x)$ la température du liquide réfrigérant. A l'instant t le granule C peut être trouvé au lieu $x = c_1 t$, et ainsi la température de son milieu au moment t sera

$$F(c_1 t) = f(t).$$

En raison de ce fait nous prescrivons la condition aux limites suivante sur la surface S du granule:

$$(a) \quad u(P, t) + \alpha \frac{\partial u(P, t)}{\partial n} = f(t) \quad (t > 0, P \in S), \quad \left(\alpha = \frac{k}{h} \geq 0 \right).$$

Nous distinguerons deux cas selon que h est fini ou bien infini. Dans le premier cas $\alpha > 0$, dans le second $\alpha = 0$.

En considérant que la fonction $f(t)$ figurant dans la condition (a) est aussi inconnue, nous avons besoin d'une équation de plus contenant la fonction $f(t)$. Cette équation sera fournie par le bilan thermique du système réfrigérant.

Considérons le segment $(x, x + dx)$ du système. Les granules arrivent à ce segment pendant la durée $t = \frac{x}{c_1}$ mesurée à partir de leur entrée dans le tuyau. Dans ce segment un granule perd la quantité de chaleur

$$-k \int_S \frac{\partial u \left(P, \frac{x}{c_1} \right)}{\partial n} dS dt$$

pendant le temps dt , donc tous les granules qui se trouvent dans le segment $(x, x + dx)$ perdent ensemble la quantité de chaleur

$$Q_1 = -k \int_S \frac{\partial u \left(P, \frac{x}{c_1} \right)}{\partial n} dS dt NK dx$$

pendant le temps dt . Le médium réfrigérant gagne dans ce segment $(x, x + dx)$ la quantité de chaleur

$$Q_2 = [F(x) - F(x + dx)] w K c_2 \rho_2 \sigma_2 dt = \\ = -w K c_2 \rho_2 \sigma_2 F'(x) dx dt$$

pendant le même temps. L'équation cherchée est fournie par l'équation $Q_1 = Q_2$ exprimant la conservation de la chaleur, l'équation nommée bilan thermique:

$$-k \int_S \frac{\partial u(P, \frac{x}{c_1})}{\partial n} dS dt N K dx = -w K c_2 \rho_2 \sigma_2 F'(x) dx dt.$$

En effectuant les simplifications possibles et en introduisant la variable $t = \frac{x}{c_1}$ au lieu de x , notre équation prend la forme:

$$(b) \quad \int_S \frac{\partial u(P, t)}{\partial n} dS = \gamma f'(t) \quad \left(\gamma = \frac{c_2}{c_1} \frac{w \rho_2 \sigma_2}{N k} > 0 \right).$$

Finalement le problème sera complètement déterminé par la température initiale du médium réfrigérant et par la répartition initiale de la température des granules:

$$(c_1) \quad F(l) = f(t_l) = F_l \quad \left(t_l = \frac{l}{c_1} \right),$$

$$(c_2) \quad u(P, 0) = u_0(P) \quad (P \in C).$$

Dans ce qui suit nous aurons besoin de la connaissance de la valeur $F(0) = F_0$. F_0 peut être déterminé à partir de la mesure de la réfrigération à atteindre par le système réfrigérant. En effet, supposons que nous voulons dériver une quantité q de chaleur de chaque granule au cours de la réfrigération. Alors nous devons dériver en somme la quantité de chaleur

$$q_1 = q K c_1 N dt$$

de la matière à réfrigérer pendant le temps dt . La quantité totale de chaleur gagnée par le médium réfrigérant pendant le temps dt est:

$$q_2 = w K c_2 \rho_2 \sigma_2 (F_0 - F_l) dt.$$

F_0 peut être calculé à partir de la condition $q_1 = q_2$:

$$F_0 = F_l + q \frac{c_1}{c_2} \frac{N}{w \rho_2 \sigma_2}.$$

De cette façon la condition (c_1) peut être remplacée par la condition

$$(c_1^*) \quad F(0) = f(0) = F_0.$$

Enfin nous résumons toutes les équations et conditions du problème:
L'équation de la chaleur:

$$(1) \quad \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} = a^2 \frac{\partial u}{\partial t} \quad (a^2 > 0).$$

La condition aux limites:

$$(a) \quad u(P, t) + a \frac{\partial u(P, t)}{\partial n} = f(t) \quad (t > 0, P \in S), (a \geq 0).$$

Le bilan thermique:

$$(b) \quad \int_S \frac{\partial u(P, t)}{\partial n} dS = \gamma f'(t) \quad (t > 0), (\gamma > 0).$$

Les conditions initiales:

$$(c_1) \quad f(t_i) = F_i \quad (t_i > 0),$$

ou se qui revient au même:

$$(c_1^*) \quad f(0) = F_0,$$

et

$$(c_2) \quad u(P, 0) = u_0(P) \quad (P \in C).$$

Nous remarquons que la méthode mentionnée au commencement du § 4 du travail [1] nous permet de ramener la solution du problème, dans le cas où $u_0(P) \neq 0$, à la solution d'un problème où $u_0(P) \equiv 0$. Mais dans ce nouveau problème la condition (b) augmente d'un membre, conformément au membre $Q(t)$ figurant dans la condition (b) de l'ouvrage [1].

Dans ce qui précède nous n'avons pas tenu compte des conditions de continuité et de dérivabilité des fonctions en question. Pour fixer les idées, nous faisons les restrictions suivantes:

Condition **F**: la fonction $u(P, t)$ admet des secondes dérivées continues selon les coordonnées de lieu x_i et une première dérivée continue selon le temps t dans le domaine $C \times (t > 0)$, elle satisfait ici à l'équation (1) de la chaleur et remplit les conditions (a) et (b) pour $t > 0$. (Ici nous avons implicitement supposé que la fonction $u(P, t)$ est dérivable même sur la surface S du domaine C pour $t > 0$.)

Condition **F** _{$a > 0$} : au cas où $a > 0$, $u(P, t)$ est continue dans le domaine $(C + S) \times (t \geq 0)$.

Condition **F** _{$a = 0$} : au cas où $a = 0$, $u(P, t)$ est continue dans le domaine $\{C \times (t \geq 0)\} + \{S \times (t > 0)\}$, et possède une discontinuité de type **A** sur la portion $S \times (t = 0)$ de ce domaine.

Condition **G**: la fonction $f(t)$ admet une dérivée continue pour $t > 0$ et est continue même au point $t = 0$.

§ 2. Principes du maximum; unicité de la solution

Dans ce qui suit nous aurons besoin de trois lemmes. Les deux premiers d'entre eux sont identiques aux deux premiers lemmes figurant dans le travail [1] de l'auteur, c'est pourquoi nous ne faisons pas connaître ici leurs démonstrations.

Lemme 1. *Soit*

$$D = C \times (0 < t \leq T),$$

$$H = \{C \times (t = 0)\} + \{S \times (0 \leq t \leq T)\},$$

de plus

$$\Sigma = S \times [(t = t_1) + \dots + (t = t_i)] \quad (0 \leq t_1 < \dots < t_i \leq T).$$

Supposons que la fonction $V(P, t)$ satisfasse à l'équation de la chaleur (1) dans D , soit continue dans $D + H - \Sigma$ et possède une discontinuité de type A sur la portion Σ de la frontière de D . Soit M la limite supérieure des valeurs de la fonction prises sur $H - \Sigma$:

$$M = \sup_{(P, t) \in H - \Sigma} V(P, t).$$

Alors, la fonction $V(P, t)$ ne peut pas prendre une valeur supérieure à M dans D .

Lemme 2. *Supposons que le point P_0 de la frontière S du domaine C possède la propriété B et que la fonction $V(P, t)$ non identiquement constante satisfait sur $C \times (0 < t \leq T)$ à l'équation (1), est continue dans*

$$B = \{C \times (0 < t \leq T)\} + (P_0, T)$$

et prend le minimum de ses valeurs prises dans B au point (P_0, T) . Alors

$$\lim_{\substack{P \rightarrow P_0 \\ P \in l}} \inf \frac{V(P, T) - V(P_0, T)}{PP_0} > 0,$$

où l est une demie-droite partant de P_0 qui forme un angle aigu avec la normale intérieure de la sphère Γ appartenant à P_0 . (Quant à Γ , voir la définition de la propriété B.)

Le fait, qu'au cas $\alpha = 0$ la fonction $u(P, t)$ ne doit pas être continue (et en effet, elle ne l'est pas) sur la partie $S \times (t = 0)$ de la frontière de l'ensemble $C \times (t \geq 0)$ (voir la condition $F_{\alpha=0}$) causera quelque difficulté dans les démonstrations des principes du maximum, c'est pourquoi nous avons besoin du lemme suivant:

Lemme 3. *Soit la surface S du domaine C une surface de Jordan composée des surfaces de Liapounoff en nombre fini, de plus supposons qu'existe la fonction calorique d'unité $E_C^{(0)}(P, t)$. Alors, au cas où $\alpha = 0$, les inégalités suivantes sont valables:*

a) *si*

$$\inf_{P \in C} u_0(P) \geq F_1,$$

alors

$$F_0 \geq F_1;$$

b) si

$$F_l \leq \sup_{P \in C} u_0(P),$$

alors

$$F_0 \leq \sup_{P \in C} u_0(P).$$

Démonstration. a) Supposons que $F_0 < F_l$. Soit $F_0 < \mu \leq F_l$, et soit t^* la valeur la plus petite ($t^* > 0$), pour laquelle $f(t^*) = \mu$ (voir fig. 1). Soit alors $u^*(P, t)$ la solution de l'équation (1) dans $C \times (0 < t \leq t^*)$ qui satisfait à la condition initiale et à la condition aux limites suivantes:

$$u^*(P, 0) = \inf_{R \in C} u_0(R) \quad (P \in C),$$

$$u^*(P, t) = \mu \quad (0 < t \leq t^*, P \in S),$$

de plus qui possède une discontinuité de type A sur la partie $S \times (t = 0)$ de la frontière de l'ensemble $C \times (0 < t \leq t^*)$.

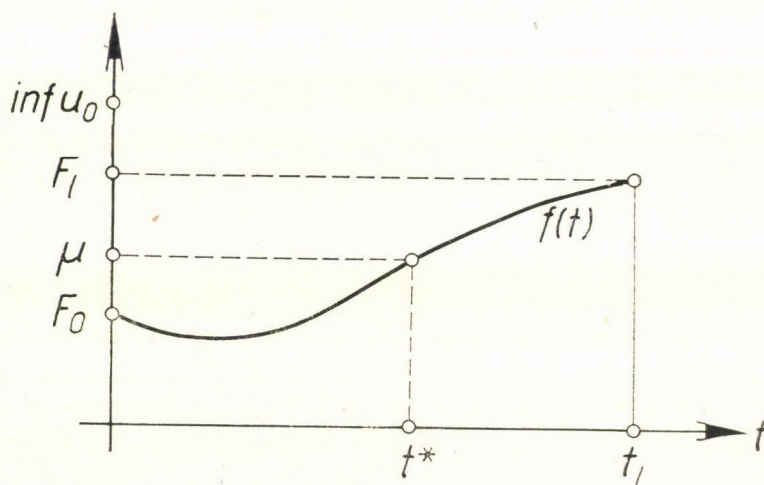


Figure 1

Alors il découle du lemme 1, que

$$(2) \quad u^*(P, t) \geq u(P, t) \quad (0 \leq t \leq t^*, P \in C).$$

D'autre part, vu que le minimum de $u^*(P, t)$ sur un ensemble arbitraire $(C + S) \times (0 < t \leq t^*)$ ($0 < t \leq t^*$) est justement la valeur μ prise sur la frontière S , il s'ensuit que

$$(3) \quad \frac{\partial u^*(P, t)}{\partial n} \leq 0 \quad (0 < t \leq t^*, P \in S).$$

En utilisant l'équation (1) et le théorème de Gauss³, on obtient le rapport suivant:

$$\begin{aligned}
 \int_C [u(P, 0) - u(P, t^*)] dC &= - \int_C \int_0^{t^*} \frac{\partial u(P, t)}{\partial t} dt dC = - \frac{1}{a^2} \int_C \int_0^{t^*} \Delta u(P, t) dt dC = \\
 (4) \quad &= - \frac{1}{a^2} \int_0^{t^*} \int_C \Delta u(P, t) dC dt = - \frac{1}{a^2} \int_0^{t^*} \int_S \frac{\partial u(P, t)}{\partial n} dS dt.
 \end{aligned}$$

En considérant l'inégalité (2), en transformant de la manière ci-dessus l'intégrale de volume qui se produit, à l'aide du théorème de Gauss, en une intégrale de surface, enfin en faisant usage de l'inégalité (3), nous arrivons à l'inégalité suivante:

$$\int_C [u(P, 0) - u(P, t^*)] dC \geq \int_C [u^*(P, 0) - u^*(P, t^*)] dC = - \frac{1}{a^2} \int_0^{t^*} \int_S \frac{\partial u^*(P, t)}{\partial n} dS dt \geq 0.$$

Après avoir comparé notre résultat à (4), nous obtenons

$$(5) \quad \int_0^{t^*} \int_S \frac{\partial u(P, t)}{\partial n} dS dt \leq 0.$$

Intégrons tous les deux membres de l'équation (b) selon t entre les bornes $t = 0$ et $t = t^*$:

$$(6) \quad \int_0^{t^*} \int_S \frac{\partial u(P, t)}{\partial n} dS dt = \gamma \int_0^{t^*} f'(t) dt = \gamma[f(t^*) - f(0)] = \gamma(\mu - F_0).$$

Or c'est impossible, parce que selon (5) le premier membre de cette équation est non-positif, cependant le second membre est une quantité certainement positive d'après l'hypothèse.

b) Dans ce cas aussi nous pouvons procéder de la même manière qu'au cas précédent a). Soient

$$\sup_{P \in C} u_0(P) \leq \mu < F_0$$

et t^* la valeur la plus petite ($t^* > 0$) pour laquelle $f(t^*) = \mu$. En définissant la fonction auxiliaire $u^*(P, t)$ de la même façon qu'au cas a), et en considérant que la valeur μ est le maximum de $u^*(P, t)$ dans l'intervalle $0 \leq t \leq t^*$, nous

³ Pour pouvoir appliquer le théorème de Gauss on a besoin des conditions supplémentaires se rapportant à la fonction $u(P, t)$ pour $t = 0$. Mais on peut voir aisément que si l'on écrit la limite $t \rightarrow +0$ au lieu de $t = 0$, l'exactitude de nos transformations sera vérifiée.

pouvons aboutir à l'inégalité (6) avec cette différence que maintenant le premier membre resp. le second membre se présentent non-négatif resp. certainement négatif. C. qu. f. d.

Après cela passons à l'étude des principes du maximum.

Pour pouvoir nous exprimer plus facilement convenons de ce que c'est en effet une réfrigération (et non le chauffage) qui s'effectue dans l'appareil réfrigérant. Alors notre premier théorème exprime que la température des granules à réfrigérer ne peut pas être plus basse que la température d'entrée F_l du médium réfrigérant à supposer que le minimum de la température des granules au moment initial ne soit pas inférieur à F_l . Le deuxième théorème exprime que la température du médium réfrigérant ne peut pas surpasser le maximum de la température initiale des granules.

Dans nos théorèmes nous supposons que la surface S possède la propriété **B**. Mais vu qu'aux cas $\alpha = 0$ nous aurons besoin du lemme 3, dans ces cas nous devons encore supposer que la surface soit une surface de Jordan composée des surfaces de Liapounoff en nombre fini.

Théorème 1. Si

$$\inf_{P \in C} u_0(P) \geq F_l,^4$$

alors

$$\min_{P \in C+S} u(P, t) \geq F_l \quad (0 < t \leq t_1)$$

Démonstration. Supposons qu'il existe un point (P_0, t_0) ($0 < t_0 \leq t_1$) tel que

$$u(P_0, t_0) = \mu < F_l.$$

1° Soit $\alpha > 0$. Alors (en vertu de la continuité de $u(P, t)$, voir la condition $F_{\alpha > 0}$) du théorème de ТЫХОНОВ (voir p. e. [10]) découle l'existence d'un point (P_1, t_1) ($0 < t_1 \leq t_0$, $P_1 \in S$), pour lequel

$$u(P_1, t_1) = \mu.$$

De la continuité de $u(P, t)$ résulte l'existence d'un t^* , le plus petit possible ($0 < t^* \leq t_1$), pour lequel on peut encore trouver un point $P^* \in S$ tel que

$$u(P^*, t^*) = \mu,$$

mais

$$u(P, t) > \mu, \quad \text{si } (P, t) \in \{(C + S) \times (0 \leq t < t^*)\}.$$

voir fig. 2). Alors il s'ensuit du lemme 2 que

$$(7) \quad \frac{\partial u(P^*, t^*)}{\partial n} < 0.$$

Il vient de la condition (a) en vertu de (7) que

$$\min_{P \in C+S} u(P, t^*) = \mu > f(t^*).$$

⁴ Au cas de $\alpha > 0$, en vertu de la condition de continuité se rapportant à $u(P, t)$ (condition $F_{\alpha > 0}$) on peut écrire „minimum” au lieu de „infimum”.

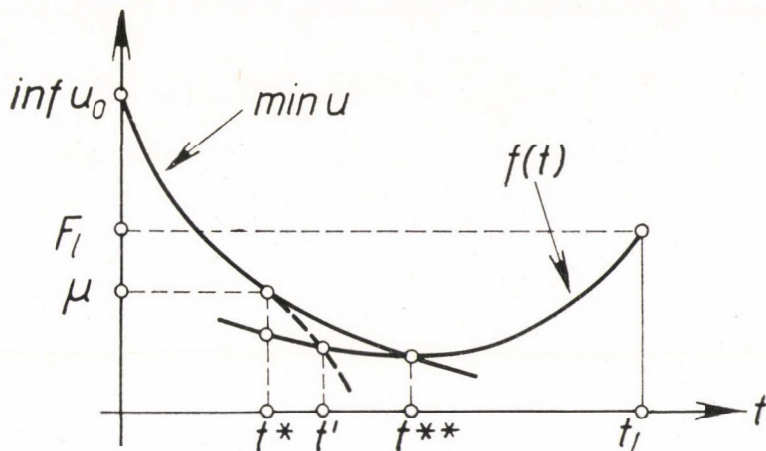


Figure 2

Pour cela à fortiori

$$u(P, t^*) > f(t^*) \quad (P \in S)$$

pour chaque point P de S , et de cette manière il découle de (a) que

$$\frac{\partial u(P, t^*)}{\partial n} < 0 \quad (P \in S)$$

pour chaque point $P \in S$.

En faisant usage de cette inégalité on obtient de (b):

$$f'(t^*) < 0.$$

Étant donné que

$$f(t^*) < F_l \text{ et } f'(t^*) < 0,$$

il doit exister un point $t^{**} > t^*$, le plus proche à droite de t^* , où

$$f'(t^{**}) = 0.$$

Dans l'intervalle $t^* < t \leq t^{**}$ la courbe $\min_{P \in C+S} u(P, t)$ ne peut pas couper la courbe $f(t)$. En effet, si elle la coupe, alors désignons par t' ($t^* < t' \leq t^{**}$) le point d'intersection le plus proche de t^* . Étant donné que la fonction $f(t)$ est strictement décroissante dans l'intervalle $t^* \leq t \leq t'$, c'est pourquoi

$$\min_{P \in C+S} u(P, t') < \min_{P \in C+S} u(P, t^*) = \min_{\substack{P \in C+S \\ 0 \leq t \leq t^*}} u(P, t),$$

et ainsi

$$\min_{P \in C+S} u(P, t') = \min_{\substack{P \in C+S \\ 0 \leq t \leq t'}} u(P, t) = u(P', t') \quad (P' \in S).$$

Donc, en vertu du lemme 2

$$\frac{\partial u(P', t')}{\partial n} < 0;$$

d'autre part en vertu de (a)

$$\frac{\partial u(P', t')}{\partial n} = 0.$$

Nous sommes ainsi arrivés à une contradiction. Par contre si les deux courbes ne se coupent pas dans l'intervalle $t^* < t \leq t^{**}$, donc même pas au point t^{**} , alors il résulte de (a):

$$\frac{\partial u(P, t^{**})}{\partial n} < 0 \quad (P \in S)$$

pour chaque point $P \in S$ et de cette manière il découle de (b):

$$f'(t^{**}) < 0,$$

ce qui contredit le choix de t^{**} .

2° Soit $\alpha = 0$. Envisageons maintenant la condition (a) réduite simplement à la condition

$$(8) \quad u(P, t) = f(t) \quad (P \in S).$$

En se basant sur le lemme 3a) on a $f(0) = F_0 \geq F_l$, et de cette façon la fonction $f(t)$ atteint le minimum de ses valeurs prises dans l'intervalle $0 \leq t \leq t_l$ à un point intérieur t^* de cet intervalle (fig. 3). (Il existe peut-être plusieurs points de cette espèce.) Il résulte de la définition de t^* et du lemme 1 que la fonction $u(P, t)$ atteint le minimum de ses valeurs prises dans $(C + S) \times (0 < t \leq t^*)$ au moment t^* , et ce qui est plus, en vertu de (8), simultanément sur toute la frontière de C . Pour cela il découle du lemme 2 que

$$\frac{\partial u(P, t^*)}{\partial n} < 0 \quad (P \in S)$$

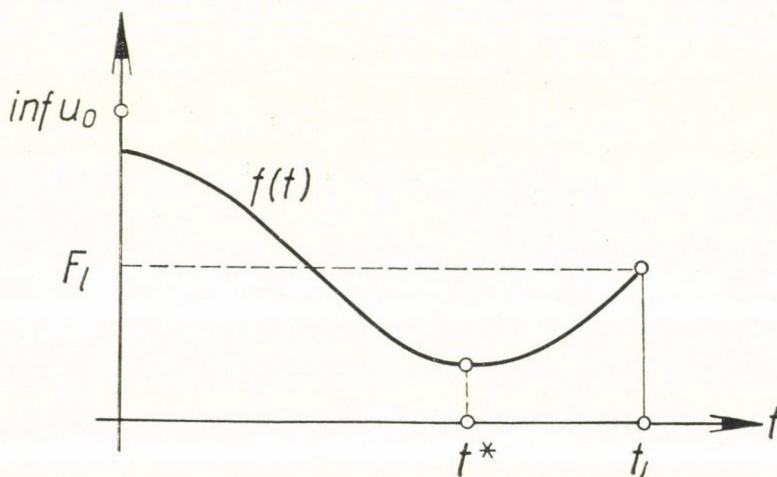


Figure 3

pour chaque point $P \in S$, et de cette manière, selon la condition (b)

$$(9) \quad f'(t^*) < 0.$$

D'autre part du fait que t^* est un point intérieur de l'intervalle $0 \leq t \leq t_l$, et ici $f(t)$ a un minimum, il résulte $f'(t^*) = 0$, ce qui contredit (9). C. qu. f. d.

Théorème 2. Si

$$F_l \leq \sup_{P \in C} u_0(P),^5$$

alors

$$f(t) \leq \sup_{P \in C} u_0(P) \quad (0 < t \leq t_l).$$

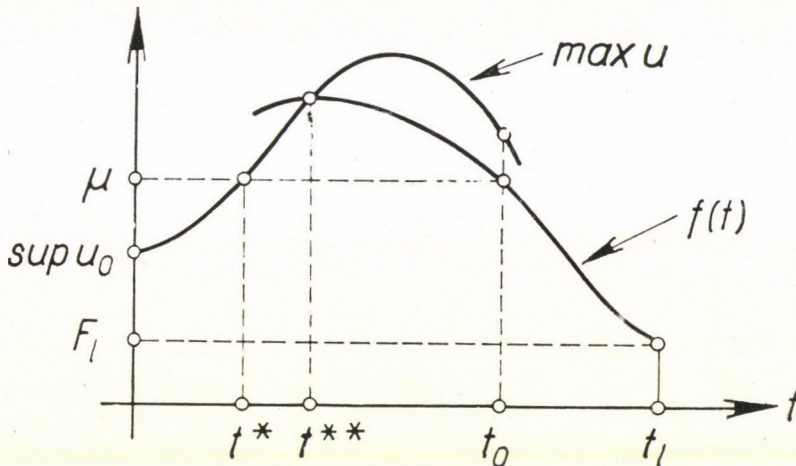


Figure 4

Démonstration. Supposons qu'il existe une valeur $t = t_0$, pour laquelle

$$f(t_0) = \mu > \sup_{P \in C} u_0(P).$$

1° Soit $\alpha > 0$. En tenant compte de ce que $f(t)$ n'est pas identiquement constante (en effet, $f(t_l) = F_l < \mu$), nous pouvons choisir les valeurs de μ et de t_0 de manière telle que

$$f'(t_0) < 0$$

(fig. 4). Alors il résulte de la condition (b) qu'il existe un point $P_0 \in S$ tel que

$$\frac{\partial u(P_0, t_0)}{\partial n} < 0,$$

et de cette manière

$$\max_{P \in C+S} u(P, t_0) > f(t_0) = \mu.$$

⁵ Au cas de $\alpha > 0$, en vertu de la condition $F_{\alpha > 0}$, on peut écrire „maximum“ au lieu de „supremum“.

Il s'ensuit de la continuité de $u(P, t)$ (condition $F_{a>0}$) en reproduisant la suite des idées suivie au commencement de la démonstration du théorème 1, qu'il existe un t^* ($0 < t^* < t_0$), le plus petit pour lequel on peut trouver un point $P^* \in S$ tel que

$$u(P^*, t^*) = \mu,$$

mais

$$u(P, t) < \mu, \quad \text{si } (P, t) \in \{(C + S) \times (0 \leq t < t^*)\}.$$

Alors il découle du lemme 2 que

$$(10) \quad \frac{\partial u(P^*, t^*)}{\partial n} > 0.$$

On obtient de la condition (a) d'après (10):

$$\max_{P \in C+S} u(P, t^*) = u(P^*, t^*) < f(t^*),$$

et de cette façon il s'ensuit de (a), conformément au résultat obtenu dans la démonstration du théorème 1, que

$$\frac{\partial u(P, t^*)}{\partial n} > 0 \quad (P \in S)$$

pour chaque point $P \in S$. De là en vertu de (b) il résulte

$$f'(t^*) > 0.$$

Étant donné que $f'(t_0) < 0$ et $f'(t^*) > 0$, il faut qu'il existe un point t^{**} ($t^* < t^{**} < t_0$), le plus proche de t^* , pour lequel

$$f'(t^{**}) = 0.$$

En reproduisant mot à mot la partie de la démonstration du théorème 1 qui vient sur ce point, n'en changeant que le terme minimum avec celui maximum et le signe $<$ avec le signe $>$, nous aboutissons à l'inégalité

$$f'(t^{**}) > 0,$$

ce qui contredit la définition de t^{**} .

2° Soit $a = 0$. Nous nous référons de nouveau au fait que la condition (a) au cas de $a = 0$ se réduit à la condition (8). Puisque selon le lemme 3b)

$$f(0) = F_0 \leq \sup_{P \in C} u_0(P),$$

la fonction $f(t)$ prend le maximum de ses valeurs prises dans l'intervalle $0 \leq t \leq t_i$ en un point t^* intérieur à cet intervalle. En suivant complètement les raisonnements employés dans le point 2° de la démonstration du théorème 1, on obtient les résultats $f'(t^*) > 0$ et $f'(t^*) = 0$, qui se contredisent mutuellement. C. qu. f. d.

Corollaire. L'unicité du système de solution $\{u(P, t), f(t)\}$ résulte d'un quelconque des principes du maximum ci-dessus de la manière bien connue.

En effet, s'il existait deux systèmes de solution, notamment $\{u_1, f_1\}$ et $\{u_2, f_2\}$, alors le système $\{u_1 - u_2, f_1 - f_2\}$ satisfairait aux conditions initiales *homogènes*, c'est pourquoi nous aurions en même temps, en vertu du théorème 1 resp. du théorème 2,

$$u_1 - u_2 \geq 0, \quad u_1 - u_2 \leq 0,$$

resp.

$$f_1 - f_2 \geq 0, \quad f_1 - f_2 \leq 0,$$

d'où l'on obtiendrait $u_1 - u_2 \equiv 0$, resp. $f_1 - f_2 \equiv 0$, et suivant la condition (a) $f_1 - f_2 \equiv 0$, resp. $u_1 - u_2 \equiv 0$. Ce dernier résultat $u_1 - u_2 \equiv 0$ découle du théorème d'unicité bien connu (voir [9] pp. 83—85) selon lequel la condition aux limites (a) (où maintenant $f(t)$ est supposée connue) détermine sans équivoque avec la condition initiale (c_2) la fonction $u(P, t)$. Pour pouvoir se référer à ce théorème d'unicité on doit se servir des restrictions imposées sur la surface S dans le lemme 3.

§ 3. La solution du problème au cas des granules de forme sphérique

Partie I

Nous nous occuperons du cas où les granules sont des sphères de rayon R . La température initiale $u_0(P)$ figurant dans la condition (c_2) soit une constante indépendante du lieu:

$$u_0(P) = u_0 \quad (= \text{constante}).$$

Alors la fonction $u(P, t)$, en ce qui concerne les coordonnées de lieu, ne dépend que de la seule coordonnée radiale r mesurée à partir du centre de la sphère:

$$u(x_1, x_2, x_3, t) = u(r, t).$$

Il s'ensuit que la condition (b) peut s'écrire sous la forme:

$$4 R^2 \pi \frac{\partial u(R, t)}{\partial r} = \gamma f'(t),$$

ou plus simplement:

$$\frac{\partial u(R, t)}{\partial r} = \delta f'(t) \quad \left(\delta = \frac{c_2}{c_1} \frac{w \varrho_2 \sigma_2}{4 R^2 \pi N k} \right).$$

Effectuons la transformation

$$u(r, t) = u_0 [1 + v(r, t)],$$

$$f(t) = u_0 [1 + g(t)].$$

Alors les nouvelles fonctions $v(r, t)$ et $g(t)$ doivent satisfaire aux équations resp. aux conditions suivantes:

$$(11) \quad \frac{\partial^2 v}{\partial r^2} + \frac{2}{r} \frac{\partial v}{\partial r} = a^2 \frac{\partial v}{\partial t} \quad (t > 0, \quad 0 \leq r < R),$$

$$(12) \quad v(R, t) + \alpha \frac{\partial v(R, t)}{\partial r} = g(t) \quad (t > 0),$$

$$(13) \quad \frac{\partial v(R, t)}{\partial r} = \delta g'(t) \quad (t < 0),$$

$$(14) \quad g(0) = g_0 = \frac{F_0}{u_0} - 1,$$

$$(15) \quad v(r, 0) = 0 \quad (0 \leq r < R).$$

Nous avons effectué cette transformation essentiellement pour obtenir une condition initiale *homogène* pour la fonction $v(r, t)$.

Nous établirons la solution du problème à l'aide du principe de Duhamel (voir [2]). Soit $E^{(a)}(r, t)$ la fonction calorique d'unité du domaine $0 \leq r < R$.

En connaissance de cette fonction nous voulons établir la solution de notre problème conformément au principe de Duhamel sous la forme

$$(16) \quad v(r, t) = \int_0^t g'(\tau) E^{(a)}(r, t - \tau) d\tau + g_0 E^{(a)}(r, t).$$

Étant donné que cette fonction satisfait à l'équation (11) et aux conditions (12) et (15) dans des conditions suffisamment faibles concernant la fonction $g(t)$ — conditions qui seront vérifiées —, notre problème se réduit à la détermination de la fonction $g(t)$ à partir de la condition (13). La fonction $g(t)$ sera calculée à l'aide de la transformation de Laplace.

Soit \mathcal{L} l'opérateur de la transformation de Laplace selon la variable t . Par la suite nous nous servirons de l'identité

$$(17) \quad \mathcal{L} \left[\int_0^t \Phi(\tau) \Psi(t - \tau) d\tau \right] = \mathcal{L}[\Phi] \mathcal{L}[\Psi]$$

se rapportant à la transformée laplacienne de l'intégrale

$$\int_0^t \Phi(\tau) \Psi(t - \tau) d\tau$$

nommée produit de convolution des fonctions $\Phi(t)$ et $\Psi(t)$.

Soient

$$\mathcal{L}[g(t)] = \varphi(p),$$

$$\mathcal{L}[E^{(a)}(r, t)] = e^{(a)}(r, p).$$

Remplaçons l'expression (16) dans la condition (13):

$$\int_0^t g'(\tau) \frac{\partial E^{(a)}(R, t - \tau)}{\partial r} d\tau + g_0 \frac{\partial E^{(a)}(R, t)}{\partial r} = \delta g'(t).$$

La transformée laplacienne de cette équation est:

$$[p\varphi(p) - g_0] \frac{\partial e^{(a)}(R, p)}{\partial r} + g_0 \frac{\partial e^{(a)}(R, p)}{\partial r} = \delta[p\varphi(p) - g_0],$$

d'où

$$\varphi(p) = \frac{\delta g_0}{p \left[\delta - \frac{\partial e^{(a)}(R, p)}{\partial r} \right]}.$$

La transformée laplacienne $e^{(a)}(r, p)$ selon [7] est:

$$e^{(a)}(r, p) = \frac{R^2 \operatorname{sh}(a \sqrt{p} r)}{rp[(R-a) \operatorname{sh}(a \sqrt{p} R) + a \sqrt{p} R \operatorname{ch}(a \sqrt{p} R)]}.$$

Cette formule est valable pour $a > 0$ ainsi que pour $a = 0$. En dérivant selon r :

$$\frac{\partial e^{(a)}(R, p)}{\partial r} = \frac{a \sqrt{p} R \operatorname{ch}(a \sqrt{p} R) - \operatorname{sh}(a \sqrt{p} R)}{p[(R-a) \operatorname{sh}(a \sqrt{p} R) + a \sqrt{p} R \operatorname{ch}(a \sqrt{p} R)]}.$$

En remplaçant cette dérivée dans l'expression de $\varphi(p)$:

$$(18) \quad \varphi(p) = \delta g_0 \frac{(R-a) \operatorname{sh}(a \sqrt{p} R) + a \sqrt{p} R \operatorname{ch}(a \sqrt{p} R)}{[\delta(R-a)p + 1] \operatorname{sh}(a \sqrt{p} R) + (a \delta p - 1) a \sqrt{p} R \operatorname{ch}(a \sqrt{p} R)}.$$

Maintenant notre problème consistera à déterminer la transformée inverse $g(t)$ de la fonction $\varphi(p)$, ce qui se fera à l'aide de la formule

$$(19) \quad g(t) = \frac{1}{2\pi i} \int_{\omega-i\infty}^{\omega+i\infty} \varphi(p) e^{pt} dp$$

de Riemann—Mellin et du théorème des résidus.

Dans la formule (19), ω signifie un nombre tel que la fonction $\varphi(p)$ est régulière dans le demi-plan $\operatorname{Re} p \geq \omega$. L'existence d'un tel ω sera démontrée dans la partie II (point 1°) du paragraphe.

Les singularités de la fonction $\varphi(p)$ ne peuvent avoir lieu qu'aux points où le dénominateur s'annule. Les racines négatives resp. positives (dans le cas où il existe des racines positives) du dénominateur, après les substitutions

$$a \sqrt{p} R = i \mu, \quad p = -\frac{\mu^2}{a^2 R^2},$$

resp.

$$a \sqrt{p} R = \kappa, \quad p = \frac{\kappa^2}{a^2 R^2},$$

sont fournies par les racines réelles différentes de zéro des équations suivantes, nommées équations caractéristiques:

$$(20.a) \quad \operatorname{tg} \mu = \frac{\mu + \frac{a \delta}{a^2 R^2} \mu^3}{1 - \frac{\delta(R-a)}{a^2 R^2} \mu^2},$$

resp.

$$(20.b) \quad th \kappa = \frac{\kappa - \frac{\alpha \delta}{a^2 R^2} \kappa^3}{1 + \frac{\delta(R - \alpha)}{a^2 R^2} \kappa^2}.$$

Nous démontrerons dans l'appendice que les racines réelles des équations (20.a) et (20.b) fournissent toutes les racines du dénominateur de la fonction $\varphi(p)$, c'est-à-dire toutes les racines du dénominateur en question sont réelles; les racines des équations (20.a) resp. (20.b) peuvent être relevées sur les figures 5.a resp. 5.b. On peut voir immédiatement sur ces figures que le dénominateur

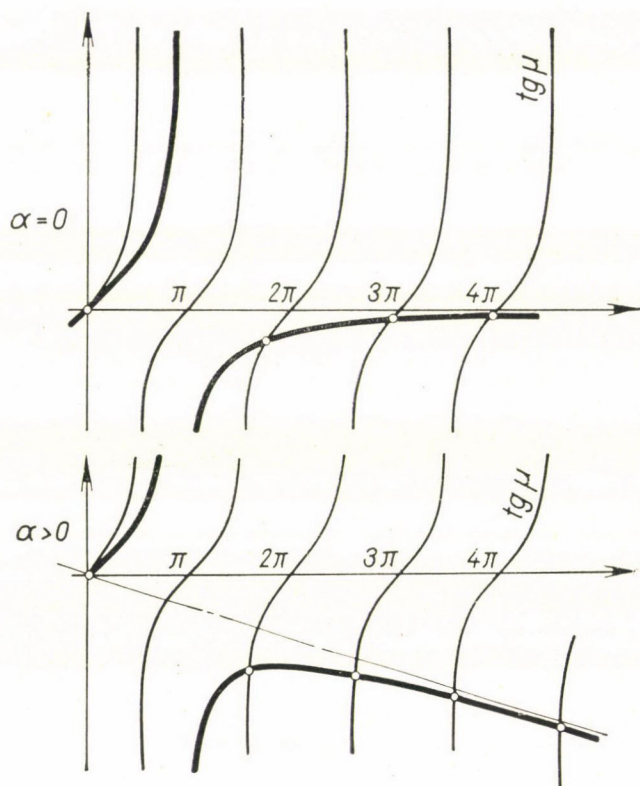


Figure 5a

de $\varphi(p)$ possède une infinité de racines négatives et un nombre fini de racines positives. (Il n'existe peut-être pas de racines positives.) Plus précisément: nous démontrerons dans l'appendice qu'au cas où

$$\frac{3 \delta}{a^2 R} \geq 1$$

le dénominateur de $\varphi(p)$ ne possède pas de racine positive, tandis qu'au cas où

$$\frac{3\delta}{a^2 R} < 1$$

il a précisément 1 racine positive. Nous reviendrons sur l'interprétation physique de ce phénomène à la fin du § 4.

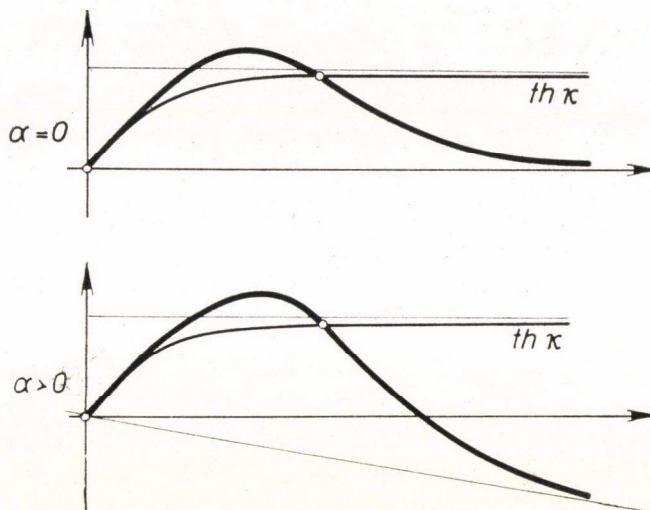


Figure 5b

Considérons la portion du chemin d'intégration figurant dans l'intégrale (19) qui se trouve entre les branches de la parabole

$$(21.a) \quad \eta^2 = \frac{4 m^2 \pi^2}{a^2 R^2} \left(\xi + \frac{m^2 \pi^2}{a^2 R^2} \right) \quad (m \text{ est un nombre entier})$$

resp. de la parabole

$$(21.b) \quad \eta^2 = \frac{4 \left(m + \frac{1}{2} \right)^2 \pi^2}{a^2 R^2} \left(\xi + \frac{\left(m + \frac{1}{2} \right)^2 \pi^2}{a^2 R^2} \right) \quad (m \text{ est un nombre entier})$$

suivant que $a > 0$ ou bien $a = 0$. Ici ξ resp. η signifient la partie réelle resp. imaginaire de la variable p : $p = \xi + i \eta$. Désignons cette portion par L_1 , et par L_2 la portion de la parabole se trouvant dans le demi-plan $\text{Re } p \leq \omega$. Comme on le verra dans la partie II (point 2°), pour des m suffisamment grands la fonction $\varphi(p)$ ne possède pas de singularité sur la courbe L_2 , et ainsi, selon le théorème des résidus:

$$(22) \quad \frac{1}{2\pi i} \int_{L_1 + L_2} \varphi(p) e^{pt} dp = \sum_{L_1 + L_2} \text{Res} [\varphi(p) e^{pt}],$$

où Σ' se rapporte aux singularités de $\varphi(p)$ se trouvant à l'intérieur de la courbe fermée $L_1 + L_2$. Nous montrerons dans la partie II (point 2°), qu'au cas où $m \rightarrow \infty$, l'intégrale prise sur L_2 tend vers zéro:

$$\lim_{m \rightarrow \infty} \int_{L_2} \varphi(p) e^{pt} dp = 0,$$

et ainsi

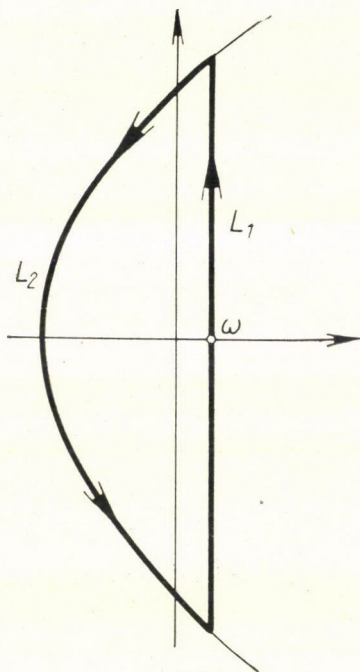


Figure 6

$$(23) \quad g(t) = \frac{1}{2\pi i} \int_{\omega - i\infty}^{\omega + i\infty} \varphi(p) e^{pt} dp = \frac{1}{2\pi i} \lim_{m \rightarrow \infty} \int_{L_1} \varphi(p) e^{pt} dp = \Sigma \text{Res}[\varphi(p) e^{pt}],$$

où Σ se rapporte à toutes les singularités de la fonction $\varphi(p)$.

Vu que la fonction à intégrer dans l'intégrale (19) prend des valeurs conjuguées aux lieux conjugués, la valeur de l'intégrale est purement imaginaire, c'est-à-dire la fonction $g(t)$ est réelle; cela est par ailleurs naturel selon la nature du problème.

Soient $\Phi(p)$ resp. $\Psi(p)$ le numérateur resp. le dénominateur de $\varphi(p)$:

$$\Phi(p) = \delta g_0[(R - a) \text{sh}(a\sqrt{p}R) + a\sqrt{p}R \text{ch}(a\sqrt{p}R)],$$

$$\Psi(p) = [\delta(R - a)p + 1] \text{sh}(a\sqrt{p}R) + (a\delta p - 1)a\sqrt{p}R \text{ch}(a\sqrt{p}R).$$

Étant donné que le point $p = 0$ est un point de ramification et une racine de $\Phi(p)$ ainsi que de $\Psi(p)$, c'est pourquoi nous calculons le résidu de la fonction $\varphi(p) e^{pt}$ au point $p = 0$ à l'aide du développement en série des fonctions $\Phi(p)$, $\Psi(p)$ et e^{pt} :

$$\begin{aligned} \frac{\Phi(p)}{a \sqrt{p} R} &= R + p \left(\frac{a^2 R^3}{6} + \frac{a a^2 R^2}{3} \right) + p^2(\dots) + \dots, \\ \frac{\Psi(p)}{a \sqrt{p} R} &= p \left(\delta R - \frac{a^2 R^2}{3} \right) + p^2 \left[-\frac{a^4 R^4}{30} + \delta \left(\frac{R}{6} + \frac{a}{3} \right) a^2 R^2 \right] + \\ &+ p^3 \left[-\frac{a^6 R^6}{840} + \delta \left(\frac{R}{120} + \frac{a}{30} \right) a^4 R^4 \right] + p^4(\dots) + \dots, \end{aligned}$$

$$e^{pt} = 1 + pt + \dots$$

Il en résulte que

$$\text{pour } \frac{a^2 R}{3 \delta} \neq 1$$

$$\text{Res}_{p=0} [\varphi(p) e^{pt}] = g_0 \frac{1}{1 - \frac{a^2 R}{3 \delta}},$$

tandis que

$$\text{pour } \frac{a^2 R}{3 \delta} = 1$$

$$\text{Res}_{p=0} [\varphi(p) e^{pt}] = g_0 \frac{10 R(R + 7 a) + 175 a^2}{7(R + 5 a)^2} + g_0 \frac{5}{\delta(R + 5 a)} t.$$

Le fait que chacun des membres des équations caractéristiques (20.a) et (20.b) est une fonction impaire, ainsi, dans la détermination des résidus, outre la racine $p = 0$, il est suffisant de nous limiter aux racines positives des équations caractéristiques. Nous désignons les racines positives de l'équation (20.a) par μ_ν ($\nu = 1, 2, 3, \dots$), et par κ la racine positive (si elle existe) de l'équation (20.b). (Comme nous l'avons déjà mentionné, l'équation (20.b) a tout au plus 1 racine positive.)

Soient p_ν ($\nu = 1, 2, 3, \dots$) resp. p_{-1} les valeurs de la variable p correspondant à $\mu = \mu_\nu$ resp. à κ . Comme nous le verrons dans l'appendice (et en partie on peut le voir sur les figures 5.a et 5.b aussi), toutes les racines positives des équations caractéristiques ont une multiplicité simple. Par conséquent

$$\text{Res}_{p=p_\nu} [\varphi(p) e^{pt}] = \frac{\Phi(p_\nu)}{\Psi'(p_\nu)} e^{p_\nu t} \quad (\nu = -1, 1, 2, 3, \dots),$$

et pour cette raison l'expression (23) peut s'écrire de la façon suivante:

$$(24) \quad g(t) = g_0(A_0^{(0)} + A_0^{(1)} t) + \sum_{\nu=-1}^{\infty} \frac{\Phi(p_\nu)}{\Psi'(p_\nu)} e^{p_\nu t},$$

où Σ^* signifie que le terme correspondant à $\nu = 0$ ne figure pas dans la somme, et

$$\text{pour } \frac{a^2 R}{3\delta} \neq 1$$

on a

$$A_0^{(0)} = \frac{1}{1 - \frac{a^2 R}{3\delta}}, \quad A_0^{(1)} = 0,$$

$$\text{pour } \frac{a^2 R}{3\delta} = 1$$

on a

$$A_0^{(0)} = \frac{10 R(R + 7a) + 175a^2}{7(R + 5a)^2}, \quad A_0^{(1)} = \frac{5}{\delta(R + 5a)}.$$

La dérivée du dénominateur $\Psi(p)$ est:

$$\begin{aligned} \Psi'(p) = & \left[\delta(R - a) + \frac{1}{2} (a\delta p - 1) a^2 R^2 \right] \text{sh}(a\sqrt{p}R) + \\ & + \left(\frac{1}{2} \delta R + a\delta \right) a\sqrt{p}R \text{ch}(a\sqrt{p}R). \end{aligned}$$

En effectuant les substitutions

$$p_\nu = -\frac{\mu_\nu^2}{a^2 R^2} \quad (\nu = 1, 2, 3, \dots)$$

et

$$p_{-1} = \frac{\kappa^2}{a^2 R^2}$$

dans (24), on obtient:

$$(25) \quad g(t) = g_0 A_0 + g_0 A_{-1} e^{\frac{\kappa^2}{a^2 R^2} t} + g_0 \sum_{\nu=1}^{\infty} A_\nu e^{-\frac{\mu_\nu^2}{a^2 R^2} t},$$

où

$$A_\nu = \frac{(R - a) \text{tg } \mu_\nu + a \mu_\nu}{\left(R - a - \frac{a^2 R^2}{2\delta} - \frac{a}{2} \mu_\nu^2 \right) \text{tg } \mu_\nu + \left(\frac{R}{2} + a \right) \mu_\nu} \quad (\nu = 1, 2, 3, \dots),$$

et

$$A_{-1} = \frac{(R - a) \text{th } \kappa + a \kappa}{\left(R - a - \frac{a^2 R^2}{2\delta} - \frac{a}{2} \kappa^2 \right) \text{th } \kappa + \left(\frac{R}{2} + a \right) \kappa}.$$

En y substituant la valeur de $\operatorname{tg} \mu_v$ de (20.a), resp. la valeur de $\operatorname{th} \kappa$ de (20.b):

$$A_v = \frac{2}{\left(3 - \frac{a^2 R}{\delta}\right) - \left(2 \frac{a}{R} + \delta \frac{R-a}{a^2 R^2}\right) \mu_v^2 - \frac{a^2 \delta}{a^2 R^3} \mu_v^4} \quad (v = 1, 2, 3, \dots),$$

resp.

$$A_{-1} = \frac{2}{\left(3 - \frac{a^2 R}{\delta}\right) + \left(2 \frac{a}{R} + \delta \frac{R-a}{a^2 R^2}\right) \kappa^2 - \frac{a^2 \delta}{a^2 R^3} \kappa^4}.$$

Considérant — comme on peut le voir de l'équation (20.a) ou bien sur la figure 5.a — que pour $a > 0$ resp. $a = 0$ on a

$$\mu_v \sim (2v+1) \frac{\pi}{2},$$

resp.

$$\mu_v \sim v\pi,$$

c'est-à-dire dans tous les deux cas

$$\frac{1}{\mu_v} = O\left(\frac{1}{v}\right),$$

on voit aisément que les coefficients A_v tendent fortement vers zéro, et ainsi la série (25) converge bien. La vitesse de convergence est augmentée par les facteurs exponentiels.

En remplaçant la forme (25) de la fonction $g(t)$ dans l'expression (16) de $v(r, t)$, nous obtenons la répartition de la température dans les granules. Sur ce point nous avons encore à démontrer que la série (25) de $g(t)$ peut être remplacée dans l'expression (16) et que la dérivation selon r dans l'expression (16) de la fonction $v(r, t)$ après le signe d'intégrale est légitime.

La série dans (25) est absolument et uniformément convergente dans l'intervalle $0 \leq t < \infty$, car au cas $a > 0$ $A_v = O\left(\frac{1}{v^4}\right)$, tandis qu'au cas $a=0$

$A_v = O\left(\frac{1}{v^2}\right)$. La série obtenue par la dérivation terme à terme de la série de $g(t)$

1. au cas $a > 0$ est absolument et uniformément convergente dans l'intervalle $0 \leq t < \infty$, car ses coefficients sont d'ordre $O\left(\frac{1}{v^2}\right)$, tandis que

2. au cas $a = 0$ pour n'importe quel $\varepsilon > 0$ elle est de nouveau absolument et uniformément convergente dans $(\varepsilon, +\infty)$, car ses coefficients sont d'ordre $O(1)$, et $g'(t) = O\left(\frac{1}{\sqrt{t}}\right)$ dans le voisinage du point $t = 0$, comme nous le démontrerons dans la partie II (point 3°).

Pour cette raison la série de la fonction $g(t)$ peut être en effet remplacée dans l'expression (16).

Comme il peut être vérifié à partir de la forme explicite de $E^{(a)}(r, t)$, dans le cas où $a = 0$, $\frac{\partial E^{(a)}(R, t)}{\partial r}$ est d'ordre $O\left(\frac{1}{\sqrt{t}}\right)$ dans le voisinage du point $t = 0$, et naturellement est bornée même dans le voisinage du point $t = 0$ pour $a > 0$; la dérivation selon r dans l'expression (16) de $v(r, t)$ peut donc être effectuée après le signe d'intégrale.

Partie II

1° Il existe une valeur réelle ω telle que le dénominateur de la fonction $\varphi(p)$ ne s'annule pas dans le demi-plan $\operatorname{Re} p \leq \omega$.

Démonstration. Considérant la relation entre p et μ , notre assertion sera vérifiée si nous montrons l'existence d'une valeur réelle Ω telle que l'équation (20.a) ne possède pas de racines dans le demi-plan $\operatorname{Im} \mu \leq \Omega$.

Étant donné que pour $\mu \rightarrow \xi - i\infty$ ($-\infty < \xi < +\infty$)

$$|\operatorname{tg} \mu| \rightarrow 1$$

uniformément en ξ , tandis qu'au cas $a > 0$ la valeur absolue du second membre de l'équation (20.a) tend vers ∞ , resp. au cas $a = 0$ elle tend vers 0 pour $\mu \rightarrow \infty$, par conséquent la valeur cherchée Ω existe.

2° Pour un m suffisamment grand le dénominateur de $\varphi(p)$ ne possède pas de singularité sur la courbe L_2 et

$$(26) \quad \lim_{m \rightarrow +\infty} \int_{L_2} \varphi(p) e^{pt} dp = 0.$$

Démonstration. Pour raison de brièveté nous nous limitons au cas $a = 0$. Le cas $a > 0$ peut être traité de la même façon.

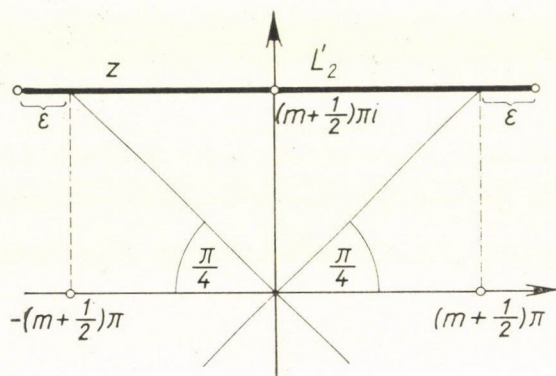


Figure 7

Employons la substitution

$$a\sqrt{p}R = z, \quad p = \frac{z^2}{a^2 R^2}$$

dans l'intégrale (26). Alors le chemin L_2 devient le chemin L'_2 donné sur la figure 7, ou $\varepsilon = O\left(\frac{1}{m}\right)$. Le résultat de la substitution est donc:

$$J = \int_{L_2} \varphi(p) e^{pt} dp = \frac{2 \delta g_0}{a^2 R} \int_{L_2'} \frac{z \operatorname{sh} z}{\left(\frac{\delta}{a^2 R} z^2 + 1 \right) \operatorname{sh} z - z \operatorname{ch} z} e^{\frac{z^2}{a^2 R^2} t} dz =$$

$$= a_1 \int_{L_2'} \frac{z}{(1 + a_2 z^2) - z \operatorname{cth} z} e^{a_3^2 z^2 t} dz,$$

où a_1, a_2 et a_3 sont des constantes.

Effectuons la substitution

$$z = \lambda + \left(m + \frac{1}{2} \right) \pi i.$$

On obtient:

$$J = a_1 \int_{-(m+\frac{1}{2})\pi-\varepsilon}^{(m+\frac{1}{2})\pi+\varepsilon} \frac{\lambda + \left(m + \frac{1}{2} \right) \pi i}{1 + a_2 \left[\lambda + \left(m + \frac{1}{2} \right) \pi i \right]^2 - \left[\lambda + \left(m + \frac{1}{2} \right) \pi i \right] \operatorname{th} \lambda} e^{a_3^2 \left[\lambda + \left(m + \frac{1}{2} \right) \pi i \right]^2 t} d\lambda.$$

Vu que $|\operatorname{th} \lambda| \leq 1$, la fraction dans la fonction à intégrer est d'ordre $O\left(\frac{1}{m}\right)$, ainsi elle ne peut pas posséder de singularité sur le segment

$$-\left(m + \frac{1}{2} \right) \pi - \varepsilon \leq \lambda \leq \left(m + \frac{1}{2} \right) \pi + \varepsilon$$

pour un m suffisamment grand. Nous obtenons l'estimation suivante pour J :

$$J = O\left(\frac{1}{m}\right) \int_{-(m+\frac{1}{2})\pi-\varepsilon}^{(m+\frac{1}{2})\pi+\varepsilon} \left| e^{a_3^2 \left[\lambda + \left(m + \frac{1}{2} \right) \pi i \right]^2 t} \right| d\lambda = O\left(\frac{1}{m}\right) \int_{-(m+\frac{1}{2})\pi-\varepsilon}^{(m+\frac{1}{2})\pi+\varepsilon} e^{a_3^2 \left[\lambda^2 - \left(m + \frac{1}{2} \right)^2 \pi^2 \right] t} d\lambda.$$

Ici l'intégrale est d'ordre $O(1)$, par conséquent

$$J = O\left(\frac{1}{m}\right).$$

3° Au cas où $a = 0$, l'ordre de grandeur de la série dérivée de la série

$$\psi(t) = \sum_{p=1}^{\infty} A_p e^{-\frac{\mu_p^2}{a^2 R^2} t}$$

est $O\left(\frac{1}{\sqrt{t}}\right)$ dans le voisinage du point $t = 0$:

$$\psi'(t) = \sum_{p=1}^{\infty} A_p' e^{-\frac{\mu_p^2}{a^2 R^2} t} = O\left(\frac{1}{\sqrt{t}}\right) \quad (t \sim 0).$$

Démonstration. Comme au cas $\alpha = 0$

$$A_p = O\left(\frac{1}{p^2}\right),$$

on obtient pour les coefficients de la série dérivée:

$$A'_p = O(1).$$

Ainsi pour des constantes convenablement choisies a_1, a_2, a_3 et M (M est un nombre entier positif), l'inégalité suivante est valable:

$$\begin{aligned} \psi'(t) &< a_1 + a_2 \sum_{p=M}^{\infty} e^{-a_3^2 p^2 t} < a_1 + a_2 \int_0^{\infty} e^{-a_3^2 v^2 t} dv = \\ &= a_1 + a_2 \frac{1}{\sqrt{t}} \int_0^{\infty} e^{-a_3^2 v^2} dv = O\left(\frac{1}{\sqrt{t}}\right). \end{aligned}$$

§ 4. Détermination de la longueur de l'appareil réfrigérant

En remplaçant l'expression (25) de la fonction $g(t)$ dans $f(t)$, puis en remplaçant cette dernière dans l'équation $F(c_1 t) = f(t)$, de plus, en considérant la relation $x = c_1 t$, on obtient la température $F(x)$ du liquide réfrigérant:

$$(27) \quad F(x) = u_0 + (F_0 - u_0) \left[A_0^{(0)} + A_0^{(1)} t + A_{-1} e^{\frac{\kappa^2}{a^2 R^2} \frac{x}{c_1}} + \sum_{p=1}^{\infty} A_p e^{-\frac{\mu_p^2}{a^2 R^2} \frac{x}{c_1}} \right].$$

La longueur l de l'appareil peut être déterminée à partir de la condition

$$(28) \quad F(l) = F_l.$$

Si $\frac{3\delta}{a^2 R} < 1$, et si nous ne tenons pas compte de la somme qui contient seulement des fonctions exponentielles de puissance négative, alors l'équation (28) peut se mettre sous la forme suivante:

$$u_0 + (F_0 - u_0) \left[A_0^{(0)} + A_{-1} e^{\frac{\kappa^2}{a^2 R^2} \frac{l}{c_1}} \right] = F_l,$$

d'où

$$l = \frac{a^2 R^2 c_1}{\kappa^2} \ln \frac{\frac{u_0 - F_l}{u_0 - F_0} - A_0^{(0)}}{A_{-1}} \quad \left(\frac{3\delta}{a^2 R} < 1 \right).$$

Le cas $\frac{3\delta}{a^2 R} = 1$, étant pratiquement sans intérêt, nous ne le traitons pas.

Si $\frac{3\delta}{a^2 R} > 1$, et si nous n'envisageons que le premier terme de la somme

dans l'expression de la fonction $F(x)$, alors l'équation (28) fournit le résultat suivant:

$$l = \frac{\alpha^2 R c_1}{\mu_1^2} \ln \frac{A_1}{\frac{u_0 - F_l}{u_0 - F_0} - A_0^{(0)}} \quad \left(\frac{3\delta}{\alpha^2 R} > 1 \right).$$

Enfin nous voudrions rendre clair la signification physique de la quantité $\frac{3\delta}{\alpha^2 R}$, et montrer pourquoi le fait que cette quantité est supérieure ou bien inférieure à 1 joue un rôle dans l'existence de la racine α .

Soit u_l la température moyenne des granules dans la section $x = l$, c'est-à-dire la température moyenne de sortie. Alors les granules transmettent au liquide réfrigérant la quantité de chaleur

$$K c_1 N \frac{4 R^3 \pi}{3} \varrho_1 \sigma_1 (u_0 - u_l)$$

par unité de temps. Le liquide réfrigérant prend pendant le même temps la quantité de chaleur

$$K c_2 w \varrho_2 \sigma_2 (F_0 - F_l)$$

des granules. Étant donné qu'il n'existe pas de transmission de chaleur vers l'extérieur, ces deux quantités sont égales:

$$K c_1 N \frac{4 R^3 \pi}{3} \varrho_1 \sigma_1 (u_0 - u_l) = K c_2 w \varrho_2 \sigma_2 (F_0 - F_l).$$

Cette équation, après avoir arrangé et utilisé les définitions de δ et de a :

$$\delta = \frac{1}{4 R^2 \pi} \frac{c_2 w \varrho_2 \sigma_2}{c_1 N k},$$

$$a^2 = \frac{\varrho_1 \sigma_1}{k},$$

peut se mettre sous la forme

$$(29) \quad \frac{3\delta}{\alpha^2 R} = \frac{u_0 - u_l}{F_0 - F_l}.$$

Examinons ce qui se passe si $l \rightarrow \infty$ pour des valeurs u_0 et F_0 fixées. Si à ce passage à la limite u_l et F_l ont des limites, alors ces limites sont égales, car sur un chemin infiniment long les granules prennent la température du liquide réfrigérant:

$$u_\infty = F_\infty = T_\infty.$$

Alors l'équation (29) peut être écrite de cette manière:

$$(29^*) \quad \frac{3\delta}{a^2 R} = \frac{u_0 - T_\infty}{F_0 - T_\infty}.$$

Au cas d'une réfrigération on a $u_0 > F_0 > T_\infty$, c'est pourquoi

$$\frac{3\delta}{a^2 R} > 1.$$

Donc, si pour $l \rightarrow \infty$ u_l et F_l possèdent des limites finies, alors

$$(30) \quad \frac{3\delta}{a^2 R} > 1;$$

il est évident que l'assertion inverse est également vraie: au cas où (30) se réalise, $u_\infty = F_\infty = T_\infty$ existe, et sa valeur peut être déterminée à partir de l'expression (27) de $F(x)$ (maintenant le membre contenant x ne se présente pas):

$$F(\infty) = u_0 + (F_0 - u_0) A_0^{(0)} = u_0 + (F_0 - u_0) \frac{1}{1 - \frac{a^2 R}{3\delta}}.$$

Ce résultat coïncide parfaitement avec le résultat (29*) obtenu par un calcul élémentaire.

Par contre, au cas

$$\frac{3\delta}{a^2 R} < 1$$

les limites de u_l et F_l n'existent pas: u_l et F_l deviennent infinis. Dans ce cas la condition (29*) donne le résultat

$$\frac{u_0 - T_\infty}{F_0 - T_\infty} \leq 1,$$

et si T_∞ existait, cela donnerait

$$u_0 \leq F_0,$$

ce qui est naturellement impossible dans le cas d'une réfrigération.

Appendice

Discussion des racines des équations caractéristiques

Nous allons démontrer le théorème suivant se rapportant au dénominateur

$$\Psi(p) = [\delta(R - a)p + 1] \operatorname{sh}(a\sqrt{p}R) + (a\delta p - 1) a\sqrt{p}R \operatorname{ch}(a\sqrt{p}R)$$

de la fonction $\varphi(p)$.

Théorème.

 a) toutes les racines de la fonction $\Psi(p)$ sont réelles,

b) les racines différentes de zéro possèdent une multiplicité simple,

 c) au cas de $\frac{3\delta}{a^2 R} \geq 1$ $\Psi(p)$ ne possède pas de racine positive, tandis qu'au

 cas de $\frac{3\delta}{a^2 R} < 1$ elle possède exactement 1 racine positive.

Démonstration. Nous effectuerons la démonstration pour $a > 0$. Au cas de $a = 0$ la démonstration peut être faite d'une façon analogue, mais à l'aide d'une discussion essentiellement plus simple.

Avec les substitutions déjà employées

$$a \sqrt{p} R = i \mu$$

resp.

$$a \sqrt{p} R = \kappa$$

de plus, en introduisant les notations

$$A = \frac{a\delta}{a^2 R^2}, \quad B = \frac{\delta}{a^2 R},$$

 l'équation $\Psi(p) = 0$, ou ce qui revient au même, les équations (20.a) resp. (20.b) peuvent s'écrire sous les formes suivantes:

$$(31) \quad \operatorname{tg} \mu = \frac{\mu + A \mu^3}{1 - (B - A) \mu^2},$$

resp.

$$(32) \quad \operatorname{th} \kappa = \frac{\kappa - A \kappa^3}{1 + (B - A) \kappa^2}.$$

Les racines réelles de $\Psi(p)$ sont fournies par les racines réelles et purement imaginaires, et seulement par celles-ci de l'équation (31).⁶ Donc l'assertion sera vérifiée si l'on montre que toutes les racines de (31) sont ou bien réelle-ou bien purement imaginaires. (Les racines purement imaginaires sont fournies par les racines réelles de l'équation (32).)

Nous écrivons l'équation (31) sous la forme

$$(31^*) \quad [1 - (B - A) \mu^2] \sin \mu - (\mu + A \mu^3) \cos \mu = 0.$$

Sur le plan de la variable complexe μ , considérons le carré Q_n qui possède comme sommets les points $n\pi(1 + i)$, $n\pi(1 - i)$, $-n\pi(1 + i)$, $-n\pi(1 - i)$

⁶ Si pour une valeur $\mu = \mu^*$ tous les deux membres de l'équation (31) deviennent infinis en même temps, alors cette valeur μ^* fournit aussi une racine de la fonction $\Psi(p)$. C'est pourquoi nous considérons ces valeurs μ^* aussi comme racines de l'équation (31). De telles valeurs μ^* ne peuvent exister qu'au cas où $B - A > 0$ et où les racines de l'équation $1 - (B - A) \mu^2 = 0$ sont en même temps des racines de l'équation $\cos \mu = 0$. Ces racines sont de multiplicité simple.

(fig. 8). Il résulte des asymptotiques valables pour $\mu \rightarrow \pm i \infty$ des fonctions $\sin \mu$ et $\cos \mu$, de plus de leurs expressions précises valables le long des droites $\operatorname{Re} \mu = \pm n \pi$, que pour un n suffisamment grand, sur la frontière du carré Q_n on a:

$$|(\mu + A \mu^3) \cos \mu| > |[1 - (B - A) \mu^2] \sin \mu|,$$

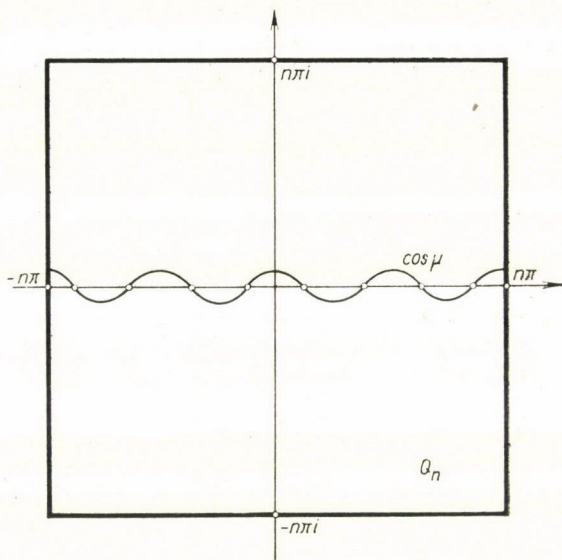


Figure 8

et ainsi selon le théorème de Rouché (voir p. e. [8]) le nombre des racines de l'équation (31*) qui se trouvent dans Q_n est égal au nombre des racines de l'équation

$$(33) \quad (\mu + A \mu^3) \cos \mu = 0$$

qui se trouvent dans Q_n (chaque racine prise autant de fois que sa multiplicité).

Pour un n suffisamment grand l'équation (33) possède $2n + 3$ racines dans Q_n . Donc l'équation (31*) et de cette manière l'équation (31) possède aussi exactement $2n + 3$ racines dans Q_n .

Dans ce qui suit nous compterons les racines de l'équation (31) dans Q_n . Pour cela nous distinguerons les 9 cas suivants:

- a) $B > \frac{1}{3}$
 - a. 1) $B - A > 0$,
 - a. 2) $B - A < 0$,
 - a. 3) $B - A = 0$,
- β) $B < \frac{1}{3}$
 - β . 1) $B - A > 0$,
 - β . 2) $B - A < 0$,
 - β . 3) $B - A = 0$,

$$\begin{aligned} \gamma) \quad B &= \frac{1}{3} \quad \gamma. 1) \quad B - A > 0, \\ & \quad \gamma. 2) \quad B - A < 0, \\ & \quad \gamma. 3) \quad B - A = 0. \end{aligned}$$

Avant de passer à la discussion détaillée des différents cas et de dessiner les figures, faisons observer avant tout les remarques suivantes:

I) En développant en série de Taylor tous les deux membres de l'équation (31) autour du point $\mu = 0$, l'équation prend la forme suivante:

$$(31) \quad \mu + \frac{1}{3} \mu^3 + \dots = \mu + B \mu^3 + \dots$$

De là on peut voir que $\mu = 0$ est une racine de l'équation (31) de multiplicité tout au moins égale à 3, tandis que

II) au cas où $B = \frac{1}{3}$, $\mu = 0$ est une racine de l'équation (31) de multiplicité tout au moins égale à 5.

III) Au cas où $B > \frac{1}{3}$ la courbe représentant le second membre de l'équation (31) en partant du point $\mu = 0$ vers la droite passe au-dessus de la courbe $\operatorname{tg} \mu$, tandis que

IV) au cas où $B < \frac{1}{3}$ elle passe au-dessous de la courbe $\operatorname{tg} \mu$.

Sur cette base on peut dessiner les figures schématiques de l'équation (31) conformément aux 9 cas distingués ci-dessus (voir fig. 9).

On peut relever les résultats suivants des figures respectives (étant donné que les figures sont symétriques par rapport à l'origine, il suffit de dessiner le demi-plan droit; nos constatations se rapportent au carré entier Q_n):

$\alpha. 1)$ Ou bien il y a de point d'intersection (p. d'int.) sur chaque branche de la courbe tangente (le cas $\alpha. 1^*$), ou bien il n'y a pas de p. d'int. sur tout au plus 2 branches (le cas $\alpha. 1^{**}$), mais alors il y a tout au moins 3 p. d'int. sur la branche principale.

$\alpha. 2)$ Il y a 3 p. d'int. sur la branche principale, il n'y a pas de p. d'int. sur les branches qui passent par les points $\mu = \pm n\pi$, et il existe des p. d'int. sur toutes les autres branches.

$\alpha. 3)$ La situation est la même qu'au cas $\alpha. 2$.

$\beta. 1)$ Il y a des p. d'int. sur toutes les branches excepté tout au plus 2 branches.

$\beta. 2)$ Il y a des p. d'int. sur toutes les branches excepté les branches passant par les points $\mu = \pm n\pi$.

$\beta. 3)$ La situation est la même qu'au cas $\beta. 2$.

$\gamma. 1)$ " " " " " " " " $\beta. 1$.

$\gamma. 2)$ " " " " " " " " $\beta. 2$.

$\gamma. 3)$ " " " " " " " " $\beta. 2$.

En tenant compte de tous ces faits, et n'oubliant pas les remarques I) és II), nous pouvons établir les faits suivants:

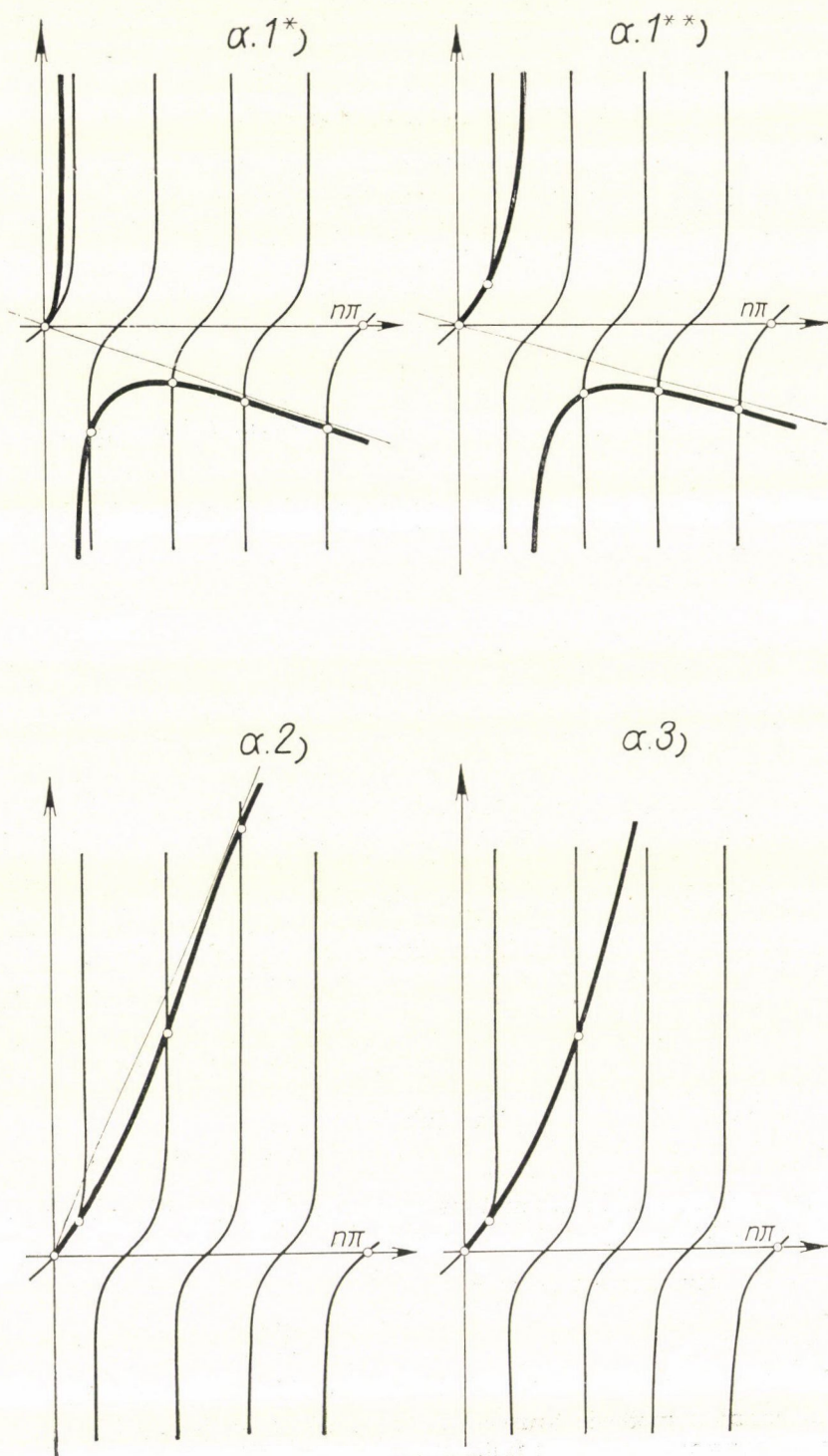


Figure 9

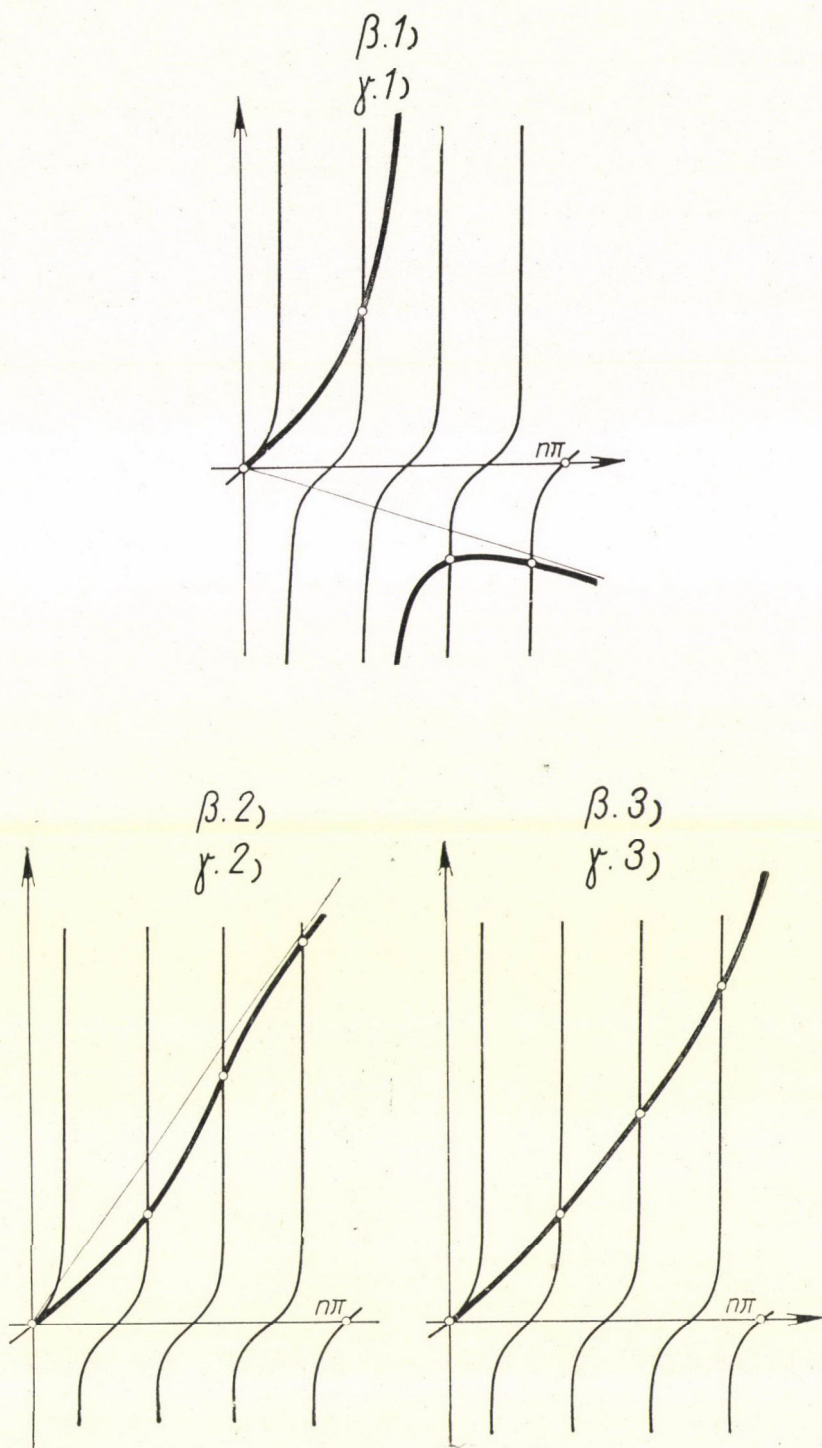


Figure 9

Au cas de $B \geq \frac{1}{3}$ l'équation (31) possède dans le carré Q_n tout au moins $2n + 3$ racines réelles. Puisque le nombre de toutes les racines dans Q_n est exactement $2n + 3$, c'est pourquoi toutes les racines de l'équation (31) sont réelles, et ces racines fournissent toutes les racines de la fonction $\Psi(p)$.

Au cas de $B < \frac{1}{3}$ l'équation (31) possède dans le carré Q_n tout au moins $2n + 1$ racines réelles. Il en résulte déjà que toutes les racines de (31) sont ou bien réelles ou bien purement imaginaires. En effet, s'il existait une racine

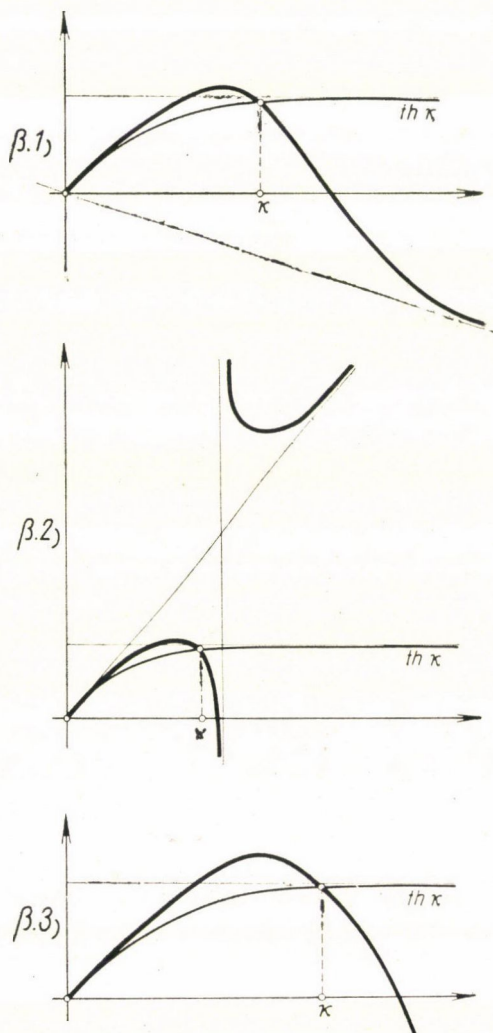


Figure 10

complexe $x + iy$ ($x \neq 0$, $y \neq 0$ réelles), alors les nombres $x - iy$, $-x + iy$, $-x - iy$ seraient aussi des racines, et de cette manière il existerait tout au moins $2n + 5$ racines dans Q_n , ce qui est impossible, parce que le nombre exact des racines dans Q_n est $2n + 3$. Comme la discussion de l'équation (32) nous le montrera, les deux racines encore manquantes sont en effet purement imaginaires, donc elles fournissent une racine positive de la fonction $\Psi(p)$.

Il nous reste encore à montrer qu'au cas où $B < \frac{1}{3}$ l'équation (32) possède pour un n suffisamment grand exactement 2 racines réelles différentes de zéro sur $(-n\pi, +n\pi)$.

En développant en série de Taylor tous les deux membres de l'équation (32) autour du point $z = 0$, l'équation prend la forme

$$(32) \quad z - \frac{1}{3}z^3 + \dots = z - Bz^3 + \dots$$

Étant donné que $B < \frac{1}{3}$, la courbe représentant le second membre de (32), en partant du point $z = 0$ vers la droite passe au-dessus de la courbe $th z$.

Nous distinguons les 3 cas suivants:

$$\beta. 1) \quad B - A > 0,$$

$$\beta. 2) \quad B - A < 0,$$

$$\beta. 3) \quad B - A = 0.$$

Conformément à ces 3 cas, les figures schématiques de l'équation (32) prennent les formes indiquées sur la fig. 10.

On peut en déduire que pour $B < \frac{1}{3}$ l'équation (32) possède toujours tout au moins 2 racines réelles différentes de zéro dans l'intervalle $(-n\pi, +n\pi)$ pour des n suffisamment grands; et en vérité elle possède dans ce cas exactement 2 racines réelles différentes de zéro, ce qui découle du nombre précis des racines dans Q_n .

Ainsi nous avons démontré toutes les trois assertions a), b) et c). C. qu. f. d.

NOTATIONS

x	coordonnée parallèle à l'axe de l'appareil réfrigérant.
x_1, x_2, x_3	coordonnées de lieu dans le système de coordonnées rattaché au granule.
P	point variable de l'espace (x_1, x_2, x_3) .
t	temps ($t = 0$ est l'instant d'entrée du granule).
l	longueur de l'appareil réfrigérant.
K	section transversale du tuyau constituant l'appareil réfrigérant.
C	granule examiné, resp. domaine occupé par le granule dans le système de coordonnées (x_1, x_2, x_3) .
S	surface du domaine C .
n	normale extérieure de la surface S .
V	volume d'un granule.

c_1, c_2	vitesse de circulation de la matière à réfrigérer resp. du médium réfrigérant.
ϱ_1, ϱ_2	densité de la matière à réfrigérer resp. du médium réfrigérant.
σ_1, σ_2	chaleur spécifique de la matière à réfrigérer resp. du médium réfrigérant.
k	conductibilité thermique de la matière à réfrigérer.
h	coefficient de transmission de chaleur entre la matière à réfrigérer et le médium réfrigérant.
N	nombre des granules par unité de volume.
$w = 1 - NV$	
$F(x)$	température du médium réfrigérant.
F_1, F_0	température d'entrée resp. de sortie du médium réfrigérant.
$u(P, t)$	température du granule.
u_0	température initiale du granule.

$$f(t) = F(c_1 t)$$

$$a^2 = \frac{\varrho_1 \sigma_1}{k}$$

$$\alpha = \frac{k}{h}$$

R rayon du granule.

$$\gamma = \frac{c_2}{c_1} \frac{w \varrho_2 \sigma_2}{N k}$$

$$\delta = \frac{\gamma}{4 R^2 \pi}$$

$$t_l = \frac{l}{c_1}$$

(Reçu le 1 Octobre 1959.)

BIBLIOGRAPHIE

- [1] ADLER, G.: „Un type nouveau des conditions aux limites de l'équation de la chaleur”. *MTA Matematikai Kutató Intézetének Közleményei* 4 (1959) 109—127.
- [2] COURANT, R. — HILBERT, D.: *Methoden der Mathematischen Physik*. Springer Verlag, Berlin, 1937. II. pp. 183—184.
- [3] FALTIN, H.: *Technische Wärmelehre* (2. Aufl.) Veb. Wilhelm-Knapp Verlag, Halle, 1953. pp. 312—320.
- [4] FRANK, PH. — MISES, R.: *Die Differential- und Integralgleichungen der Mechanik und Physik* (2. Aufl.) Verlag von Friedr. Vieweg & Sohn Akt.-Ges., Braunschweig, 1935. p. 590.
- [5] ГЮНТЕР, Н. М.: *Теория потенциала и ее применение к основным задачам математической физики*. Гостехиздат, Москва, 1953. pp. 27—28.
- [6] DE LA VALLÉE POUSSIN, CH. J.: *Cours d'Analyse Infinitésimale*. Gauthier-Villars, Paris, 1923. pp. 359—363.
- [7] ЛЫКОВ, А. В.: *Теория теплопроводности*. Гостехиздат, Москва, 1952. pp. 175—179.

- [8] SAKS, S.—ZYGmund, A.: *Analytic Functions*. Nakladem Polskiego Towarzystwa Matematycznego, Warszawa—Wrocław, 1952. p. 157.
- [9] SOMMERFELD, A.: *Partielle Differentialgleichungen der Physik*. (Vorlesungen über theoretische Physik, Band VI) Akademische Verlagsgesellschaft Geest & Portig K.—G., Leipzig, 1954. pp. 28., 78.
- [10] ТИХОНОВ, А. Н., САМАРСКИЙ, А. А.: *Уравнения математической физики*. Гостехиздат, Москва—Ленинград, 1951. pp. 184—187.

TABLE DES MATIÈRES

Introduction	327
§ 1. Établissement du problème	329
§ 2. Principes du maximum; unicité de la solution	333
§ 3. La solution du problème au cas des granules de forme sphérique Partie I	341
Partie II	350
§ 4. Détermination de la longueur de l'appareil réfrigérant	352
Appendice	354
NOTATIONS	361
BIBLIOGRAPHIE	362

SZEMCSÉS ANYAGOK ELLENÁRAMÚ HÚTÉSE

ADLER Gy.

Kivonat

Folyékony anyagnak folyadék által történő ellenáramú hűtésénél csupán a két közeg között fennálló hőátadási folyamatot szokás tekintetbe venni. Ez általában elegendő, és legtöbbször egyéb jelenségeket nem is lehet figyelemmel kísérni. Ezzel szemben abban az esetben, ha a hűtendő anyag szemcsés szerkezetű, akkor figyelemmel kell lenni az egyes szemcséken belül végbemenő hővezetési folyamatra is. A dolgozat ezen utóbbi esettel foglalkozik.

Az 1. §-ban a probléma matematikai megfogalmazását adjuk meg. Az ellenáramú hűtőkészülék legyen az x tengely $(0, l)$ szakasza mentén elhelyezett cső, melynek $x = 0$ végén a hűtendő anyag szemcséi, $x = l$ végén pedig a hűtőfolyadék áramlik be a csőbe. A csőben a szemcsék, illetve a hűtőfolyadék egymással szemben áramlanak. A stacionárius eset tárgyalására szorítkozunk. A hűtőfolyadék $F(x)$ hőmérséklete csak az x koordinátától függjön. Tegyük fel, hogy a szemcsék egybevágóak. Egy kiválasztott és a későbbiek során figyelemmel kísért reprezentáns C szemcse hőmérséklete a hozzárögzített (x_1, x_2, x_3) derékszögű koordinátarendszerben legyen $u(P, t)$, ahol P az (x_1, x_2, x_3) tér változó pontját és t a szemcsének a csőbe való belépésének pillanatától számított időt jelenti.

Bevezetve az $f(t) = F(c_1 t)$ jelölést, feladatunk a hővezetés (1) egyenletének az (a) és (b) perem- és a (c_1) (vagy ami lényegében ugyanaz, a (c_1^*)) és a (c_2) kezdeti feltételek melletti megoldását jelenti.

Ez a feladat teljesen analóg a szerző [1] dolgozatában tárgyalt problémával, és így a szemcsés anyagok ellenáramú hűtésének problémája ezen [1] dolgozatban tárgyalt újtípusú peremértékproblémára vonatkozó példának tekinthető.

A 2. §-ban a következő két maximum-elvet bizonyítjuk be:

1. Tétel. Ha

$$\inf_{P \in C} u_0(P) \geq F_l,$$

akkor

$$\min_{P \in C+S} u(P, t) \geq F_l \quad (0 < t \leq t_l).$$

2. Tétel. Ha

$$F_l \leq \sup_{P \in C} u_0(P),$$

akkor

$$f(t) \leq \sup_{P \in C} u_0(P) \quad (0 < t \leq t_l).$$

A feladat megoldásának unicitása e két maximum-elv bármelyikének közvetlen következménye.

A 3. §-ban konkrét esetre, gömb alakú szemcsék esetére vonatkozóan végezzük el a számításokat. A megoldást az

$$f(t) = u_0 [1 + g(t)]$$

helyettesítés után a (25) alatti, jól kezelhető végtelen sor állítja elő.

A 4. §-ban a hűtőberendezés hosszának meghatározására szolgáló módszer kerül ismertetésre.

Végül a függelékben a (20.a) és (20.b) alatti, ún. karakterisztikus egyenletek gyökeinek a komplex számsíkon való eloszlását vizsgáljuk meg.

ОХЛАЖДЕНИЕ ЗЕРНИСТЫХ ВЕЩЕСТВ ВСТРЕЧНЫМ ПОТОКОМ

G.Y. ADLER

Резюме

При охлаждении жидкости встречным потоком другой жидкости принимают во внимание лишь процесс теплопередачи между двумя средами. Этого обычно достаточно, в большинстве случаев невозможно следить за другими явлениями. Напротив, в том случае, когда охлаждаемое вещество имеет зернистое строение, следует принимать во внимание и процесс теплопроводности, протекающий в отдельных зернах. Работа занимается этим случаем.

В § 1 даётся математическая формулировка проблемы. Пусть охлаждающая установка есть труба, расположенная вдоль отрезка $(0, l)$ оси x , в конце $x = 0$ которой входят зерна охлаждаемого вещества, а в конце $x = l$ охлаждающая жидкость. В трубе зернистое вещество и жидкость текут во встречном направлении. Мы ограничиваемся стационарным случаем. Температура $F(x)$ охлаждающей жидкости пусть зависит лишь от координаты x . Предположим, что зёрна конгруэнтны. Рассмотрим некоторое зерно C . Пусть $u(P, t)$ обозначает его температуру в фиксированной к нему прямоугольной системе координат (x_1, x_2, x_3) . Здесь P обозначает изменяющуюся точку пространства (x_1, x_2, x_3) , а t — время измеряемое с момента входа зерна в трубу.

Введём обозначение $f(t) = F(c_1, t)$. Наша задача заключается в решении уравнения теплопроводности (1) при граничных условиях (а) и (б) и начальных условиях (c_1) (или, что в сущности одно и то же (c_1^*) и (c_2)).

Эта задача совершенно аналогична проблеме, изученной автором в работе [1]. Поэтому проблема охлаждения зернистого вещества встречным потоком может рассматриваться как пример граничной задачи, рассматриваемой в работе [1].

В § 2 доказывается два следующих принципа максимума:

Теорема 1. Если

$$\inf_{P \in C} u_0(P) \geq F_l,$$

то

$$\min_{P \in C+S} u(P, t) \geq F_l. \quad (0 < t \leq t_l).$$

Теорема 2. Если

$$F_l \leq \sup_{P \in C} u_0(P),$$

то

$$f(t) \leq \sup_{P \in C} u_0(P) \quad (0 < t \leq t_l).$$

Единственность решения задачи является непосредственным следствием любого из этих двух принципов максимума.

В § 3 вычисления проводятся для конкретного случая, случая зерен шарообразной формы. Решение после постановки

$$f(t) = u_0 [1 + g(t)]$$

представляется бесконечным рядом (25).

В § 4 описывается метод определения длины охлаждающей установки.

Наконец в приложении изучается расположение корней так называемого характеристического уравнения (20. а) и (20. б) на комплексной плоскости.

UNA APPLICAZIONE DEL CALCOLO DEGLI OPERATORI DI MIKUSIŃSKI PER LA RISOLUZIONE D'UNA EQUAZIONE ALLE DERIVATE PARZIALI

G. ADLER e G. FREUD

In questo articolo vorremmo indicare, in relazione ad un problema concreto, certi vantaggi del calcolo degli operatori di Mikusiński rispetto a quelli della trasformazione di Laplace. — Nel caso che vogliamo considerare la trasformazione di Laplace non è applicabile perchè una delle espressioni che si presentano nel corso del calcolo non è la trasformata di Laplace di nessuna funzione. Se si prescinde da questo fatto, considerato che le operazioni effettuate su questa espressione che non è una trasformata laplaciana, danno luogo soltanto a funzioni che sono trasformate laplaciane, anche con l'uso della trasformazione di Laplace riusciremmo ad ottenere un risultato giusto. Ma la giustezza di questo risultato può essere solamente verificato con l'aiuto del calcolo degli operatori di Mikusiński [3].

Il problema consiste nel cercare la soluzione dell'equazione

$$(1) \quad \frac{\partial^2 Z}{\partial x^2} + \frac{1}{x} \frac{\partial Z}{\partial x} - \frac{\partial Z}{\partial t} - \frac{1}{x^2} Z = 0$$

con la condizione iniziale e condizione al contorno seguenti:

$$(2) \quad Z(x, 0) = 0 \quad (x > 1),$$

$$(3) \quad Z(1, t) = g(t) \quad (t > 0),$$

dove $g(t)$ è una funzione con una derivata continua per intervalli.

Prima trattiamo il caso speciale

$$g(t) \equiv 1 \quad (t > 0),$$

e la soluzione corrispondente sarà indicata con $z(x, t)$.

Si può verificare che la funzione

$$z_p = \frac{1}{x} e^{-\frac{x^2}{4t}}$$

è una soluzione particolare dell'equazione (1). Per questo cerchiamo la soluzione del nostro problema nella forma

$$(4) \quad \{z\} = \{f\} * \left\{ \frac{1}{x} e^{-\frac{x^2}{4t}} \right\} = \frac{1}{x} \{f\} * \left\{ e^{-\frac{x^2}{4t}} \right\}^1$$

¹ Qui $\{ \}$ significa un operatore. Se $f(t)$ è una funzione, allora $\{f\}$ indica la operatore corrispondente secondo Mikusiński [3]. L'asterisco $*$ è il segno della moltiplicazione di operatori.

Se sappiamo che l'operatore $\{f\}$ in (4) corrisponde ad una funzione $f(t)$, allora l'espressione (4) identica col prodotto di convoluzione

$$(5) \quad z = \frac{1}{x} \int_0^t f(\tau) e^{-\frac{x^2}{4(t-\tau)}} d\tau,$$

e così potremmo applicare la trasformazione di Laplace. Ma poichè l'operatore $\{f\}$, come vedremo, non corrisponde a nessuna funzione, la trasformazione di Laplace non è applicabile. Malgrado questo l'operatore $\{z\}$ stesso corrisponderà ad una funzione.

Per poter costruire la soluzione dell'equazione (1) nella forma (4), dobbiamo verificare che

1° l'operatore (4) possiede una seconda derivata continua secondo il parametro x ,

2° l'equazione (1) è soddisfatta da (4).

Considerato che la funzione z_p soddisfa all'equazione (1), la dimostrazione di 2° si riduce alla dimostrazione del fatto che la condizione (2) è verificato dal (4).

L'asserzione 1° è una conseguenza immediata della definizione della derivata continua secondo un parametro; infatti, sotto (4) l'operatore $\{z\}$ è prodotto in forma d'un prodotto di operatori d'un operatore è d'una funzione di operatori che possiede una seconda derivata continua.

L'asserzione 2° non è una conseguenza a priori della forma (4), per questo deve essere dimostrata a posteriori. La dimostrazione sarà effettuata alla fine dell'articolo.

Sostituiamo (4) nella condizione (3):

$$\{f\} * \left\{ e^{-\frac{1}{4t}} \right\} = \{1\}.$$

Quindi

$$\{f\} = \frac{1}{\left\{ e^{-\frac{1}{4t}} \right\}}.$$

(Da qui si può vedere che l'operatore $\{f\}$ infatti non è una funzione; la divisione significa la divisione di operatori e non la divisione dei valori di funzione.) Sostituendo questa forma di $\{f\}$ in (4), otteniamo la soluzione del nostro problema nella forma di operatore

$$\{z\} = \frac{1}{x} \left\{ \frac{1}{e^{-\frac{1}{4t}}} \right\} * \left\{ e^{-\frac{x^2}{4t}} \right\},$$

o riordinato:

$$\{z\} = \frac{1}{x} \{1\} * \left\{ \begin{matrix} e^{-\frac{x^2}{4t}} \\ e^{-\frac{1}{4t}} \end{matrix} \right\}.$$

Poichè

$$\{1\} = \frac{1}{s},$$

$$\left\{e^{-\frac{x^2}{4t}}\right\} = \frac{x}{\sqrt{s}} K_1(x\sqrt{s}),$$

dove K_n è la funzione modificata di Bessel ([2] p. 29), allora $\{z\}$ può essere scritta nella forma

$$\{z\} = \frac{1}{x} \frac{1}{s} \frac{\frac{x}{\sqrt{s}} K_1(x\sqrt{s})}{\frac{1}{\sqrt{s}} K_1(\sqrt{s})} = \frac{1}{s} \frac{K_1(x\sqrt{s})}{K_1(s)}.$$

Se verifichiamo che nella precedente espressione di $\{z\}$

$$(6) \quad F(s) = \frac{1}{x} \frac{K_1(x\sqrt{s})}{K_1(s)}$$

esiste come una trasformata di Laplace, allora risulta ([3] p. 322) che $\{z\}$ corrisponde ad una funzione $z(t)$ che può essere scritta nella forma

$$(7) \quad z(t) = \mathcal{L}^{-1} \left[\frac{1}{s} \frac{K_1(x\sqrt{s})}{K_1(\sqrt{s})} \right],$$

dove \mathcal{L} significa la trasformazione di Laplace.

Essendo

$$K_1(y) \sim \sqrt{\frac{\pi}{2y}} e^{-y}$$

uniformemente per

$$|\arg y| < \frac{\pi}{2}$$

([4] p. 202), l'assintotica della funzione $F(s)$ è:

$$F(s) \sim \frac{1}{\sqrt{x}} \frac{1}{s} e^{-\sqrt{x}(s-1)} \quad (x > 1)$$

Da qui si può facilmente verificare che si realizza per un valore qualunque $\sigma \geq 0$ che $F(s) \rightarrow 0$ nel semi-piano $\operatorname{Re} s \geq \sigma$ in senso due-dimensionale se $s \rightarrow \infty$, di più, che l'integrale

$$\int_{-\infty}^{+\infty} |F(\xi + i\eta)| d\eta \quad (\xi \geq 0)$$

è convergente. Così si può applicare il teorema 3 che si trova al p. 263 del [1], in virtù di cui $F(s)$ è infatti la trasformata di Laplace d'una funzione.

Così risulta che la produzione sotto (7) di z è possibile.

L'inversione della trasformazione di Laplace sarà eseguita con l'aiuto della formola di Riemann—Mellin. Secondo questo

$$(8) \quad z = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \frac{1}{s} \frac{K_1(x\sqrt{s})}{K_1(\sqrt{s})} e^{st} ds,$$

dove σ è un valore reale tale che l'integrando non possiede singolarità nel semipiano $\operatorname{Re} s \geq \sigma$ (fig. 1).

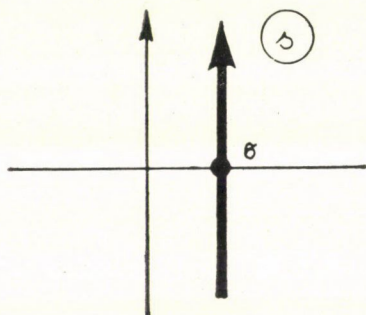


Figura 1

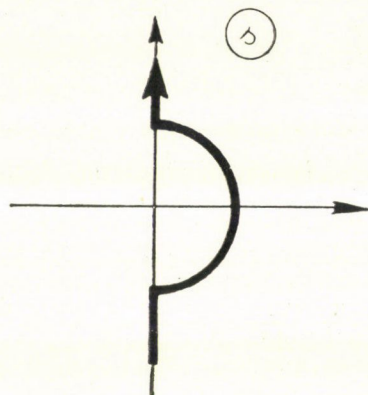


Figura 2

Poichè la funzione $K_1(s)$ non ha né radice né singolarità nel dominio $-\pi \leq \arg s \leq \pi$, per questo la sola singolarità dell'integrando si trova nel punto $s = 0$, e così il cammino d'integrazione può essere deformato nella forma data in fig. 2. Siano R_1 ed R_2 gli archi

$$R_1: \quad |s| = R, \quad \frac{\pi}{2} \leq \arg s \leq \pi,$$

$$R_2: \quad |s| = R, \quad \pi \leq \arg s \leq \frac{3\pi}{2}.$$

Si può dimostrare che

$$\lim_{R \rightarrow \infty} \int_{R_i} \frac{1}{s} \frac{K_1(x\sqrt{s})}{K_1(\sqrt{s})} e^{st} ds = 0 \quad (i = 1, 2),$$

e così il cammino d'integrazione in fig. 2 può essere deformato di nuovo ad un circolo col centro nell'origine ed al margine superiore ed inferiore dell'asse reale negativo (fig. 3).

Applicando la sostituzione $s = p^2$ si ottiene per la funzione z la forma seguente:

$$z = \frac{1}{\pi i} \int_L \frac{1}{p} \frac{K_1(xp)}{K_1(p)} e^{p^2 t} dp,$$

dove L è il nuovo cammino d'integrazione nel piano della variable p (fig. 4). Sia δ il raggio della semi-circonferenza (col centro nell'origine), allora questo ultimo integrale può essere scritto più dettagliatamente così:

$$z = \frac{1}{\pi i} \left[\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \frac{K_1(x\delta e^{i\varphi})}{K_1(\delta e^{i\varphi})} e^{\delta^2 e^{2i\varphi} t} i d\varphi + \int_{-\infty}^{-\delta} \frac{1}{r} \frac{K_1(xir)}{K_1(ir)} e^{-r^2 t} dr + \int_{+\delta}^{+\infty} \frac{1}{r} \frac{K_1(xir)}{K_1(ir)} e^{-r^2 t} dr \right].$$

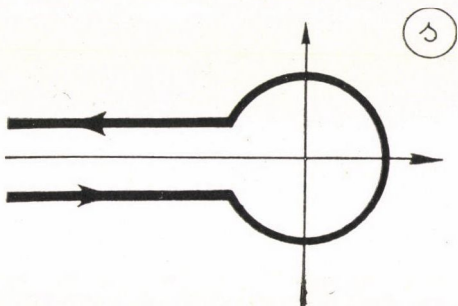


Figura 3

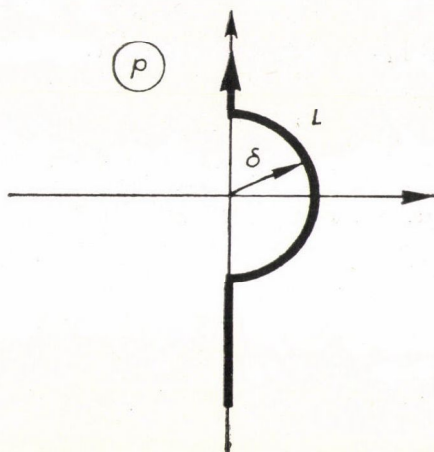


Figura 4

Poichè δ può essere arbitrariamente piccolo e come vedremo il limite del secondo membro dell'equazione esiste per $\delta \rightarrow 0$, possiamo formare il limite per $\delta \rightarrow 0$:

$$z = \frac{1}{\pi i} \left[\lim_{\delta \rightarrow 0} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \frac{K_1(x\delta e^{i\varphi})}{K_1(\delta e^{i\varphi})} e^{\delta^2 e^{2i\varphi} t} i d\varphi + \lim_{\delta \rightarrow 0} \left(\int_{-\infty}^{-\delta} \frac{1}{r} \frac{K_1(xir)}{K_1(ir)} e^{-r^2 t} dr + \int_{+\delta}^{+\infty} \frac{1}{r} \frac{K_1(xir)}{K_1(ir)} e^{-r^2 t} dr \right) \right].$$

Giacchè nel caso $y \rightarrow 0$

$$K_1(y) \sim \frac{1}{y},$$

il primo limite può essere calcolato immediatamente:

$$\lim_{\delta \rightarrow 0} \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \frac{K_1(x\delta e^{i\varphi})}{K_1(\delta e^{i\varphi})} e^{\delta^2 e^{2i\varphi} t} i d\varphi = \frac{\pi i}{x}.$$

I due integrali del secondo limite, sostituendo $(-r)$ in luogo di r nel primo integrale, possono essere scritti come un singolo integrale:

$$\int_{-\infty}^{-\delta} + \int_{+\delta}^{+\infty} = \int_{+\delta}^{+\infty} \frac{1}{r} \left[\frac{K_1(xir)}{K_1(ir)} - \frac{K_1(-xir)}{K_1(-ir)} \right] e^{-r^2 t} dr.$$

Formato il limite per $\delta \rightarrow 0$ e considerata la relazione

$$K_1(-x) = -K_1(x) + i\pi I_1(x)$$

([2] p. 29), z può essere scritto nella forma seguente:

$$z = \frac{1}{x} + \frac{1}{\pi i} \int_0^{+\infty} \frac{1}{r} \left[\frac{K_1(xir)}{K_1(ir)} - \frac{-K_1(xir) + i\pi I_1(xir)}{-K_1(ir) + i\pi I_1(ir)} \right] e^{-r^2 t} dr.$$

Utilizzando le relazioni

$$K_1(iy) = \frac{\pi}{2} [-J_1(y) + iN_1(y)],$$

$$I_1(iy) = iJ_1(y),$$

([2]) pp. 26, 28), z ottiene finalmente la forma definitiva:

$$z = \frac{1}{x} + \frac{2}{\pi} \int_0^{+\infty} \frac{1}{r} \frac{N_1(r) J_1(xr) - N_1(xr) J_1(r)}{J_1^2(r) + N_1^2(r)} e^{-r^2 t} dr.$$

(Si può verificare immediatamente che questa funzione soddisfa all'equazione (1) perchè anche le funzioni

$$J_1(xr) e^{-r^2 t}, \quad N_1(xr) e^{-r^2 t}$$

soddisfanno a quella. Inoltre si vede colla sostituzione $x = 1$, quando il numeratore dell'integrando sarà identicamente zero, che la funzione soddisfa anche alla condizione (3).)

Finalmente abbiamo ancora a dimostrare che la soluzione ottenuta verifica la condizione (2). Lo dimostreremo in base alla forma (8) della funzione z .

In virtù dell'assintotica della funzione $F(s)$, l'integrale sotto (8) è uniformemente convergente in t ($0 \leq t \leq \mu$) per $x > 1$ e per un μ qualunque positivo; così la transizione alla limite $t \rightarrow 0$ può essere effettuata dietro il segno dell'integrale:

$$z(x, 0) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} F(s) ds.$$

Servendoci dell'uniformità dell'assintotica di $F(s)$ per $x > 1$, inoltre della disuguaglianza relativa a $s = \sigma + i\tau$:

$$Re \sqrt{s} \geq \frac{1}{4} (\sqrt{|\sigma|} + \sqrt{|\tau|}) \quad (\sigma > 0),$$

si ottiene per $\sigma \rightarrow +\infty$:

$$|z(x, 0)| < c \frac{1}{\sqrt{x}} \int_{\sigma-i\infty}^{\sigma+i\infty} \left| \frac{1}{s} e^{-\sqrt{s}(x-1)} \right| |ds| \leq c \frac{1}{\sqrt{x}} \int_{-\infty}^{+\infty} \left| \frac{1}{\sqrt{\sigma^2 + \tau^2}} \right| e^{-\frac{|\sigma + i\tau|}{4}(x-1)} d\tau =$$

$$= c \frac{1}{\sqrt{x}} e^{-\frac{\sqrt{\sigma}}{4}(x-1)} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{\sigma^2 + \tau^2}} e^{-\frac{\sqrt{\tau}}{4}(x-1)} d\tau = 0,$$

dove c è una costante. Dunque la condizione (2) è soddisfatta da z :

$$z(x, 0) = 0.$$

Avendo costruita la soluzione per il caso speciale $g(t) \equiv 1$, la soluzione del problema generale è fornita dal principio di Duhamel:

$$Z(x, t) = g(0) z(x, t) + g' * z(x, t).$$

La soluzione può essere costruita anche nel caso quando g è un operatore che non corrisponde necessariamente ad una funzione. In questo caso la soluzione può essere scritta nella forma

$$Z(x, t) = s * g * z.$$

BIBLIOGRAFIA

- [1] DOETSCH, G.: *Handbuch der Laplace-transformation*, I. Verlag Birkhäuser, Basel, 1950.
- [2] MAGNUS, W.—OBERHETTINGER, F.: *Formeln und Sätze für die speziellen Funktionen der mathematischen Physik*. Springer-Verlag, Berlin, 1948.
- [3] MIKUSIŃSKI, J.: *Operatorenrechnung*. Deutscher Verlag der Wissenschaften, Berlin, 1957.
- [4] WATSON, G. N.: *Theory of Bessel Functions*. University Press, Cambridge, 1952.

A MIKUSIŃSKI-FÉLE OPERÁTOR SZÁMÍTÁS ALKALMAZÁSA EGY PARCIÁLIS DIFFERENCIÁLEGYENLET MEGOLDÁSÁRA

ADLER Gy. és FREUD G.

Kivonat

A szerzők az

$$(1) \quad \frac{\partial^2 Z}{\partial x^2} + \frac{1}{x} \frac{\partial Z}{\partial x} - \frac{\partial Z}{\partial t} - \frac{1}{x^2} Z = 0$$

parciális differenciálegyenlet

$$(2) \quad Z(x, 0) = 0 \quad (x > 1)$$

és

$$(3) \quad Z(1, t) = g(t) \quad (t > 0)$$

kezdeti-, illetve peremfeltételeket kielégítő megoldását (ahol $g(t)$ szakaszonként folytonosan differenciálható függvény) állítják elő a Mikusiński-féle operátorszámítás segítségével. Először a $g(t) \equiv 1$ esethez tartozó $z(x, t)$ megoldást

$$z(x, t) = \frac{1}{x} \int_0^t f(\tau) e^{-\frac{x^2}{4(t-\tau)}} d\tau$$

alakban határozzák meg. Ez a feladat egy, az $f(t)$ ismeretlenre vonatkozó konvolúció-típusú integrálegyenletre vezet, melynek a függvények körében nincs megoldása. Ezzel szemben ez az integrálegyenlet a Mikusiński-féle operátorok körében megoldható, és az így nyert f operátort $z(x, t)$ fenti kifejezésébe helyettesítve $z(x, t)$ már függvénynek adódik.

Az általános probléma megoldása $z(x, t)$ segítségével a következőképp írható fel:

$$Z(x, t) = \frac{\partial}{\partial t} \int_0^t g(\tau) z(x, t - \tau) d\tau.$$

Abban az esetben, midőn g olyan operátor, amely nem függvénynek felel meg, a megoldás a következő alakban írható fel:

$$Z(x, t) = s * g * z.$$

ПРИМЕНЕНИЕ ОПЕРАЦИОННОГО ИСЧИСЛЕНИЯ ОТ МИКУСИНСКИ К РЕШЕНИЮ ОДНОГО УРАВНЕНИЯ В ЧАСТНЫХ ПРОИЗВОДНЫХ

GY. ADLER и G. FREUD

Резюме

Авторы представляют решение дифференциального уравнения

$$(1) \quad \frac{\partial^2 Z}{\partial x^2} + \frac{1}{x} \frac{\partial Z}{\partial x} - \frac{\partial Z}{\partial t} - \frac{1}{x^2} Z = 0,$$

удовлетворяющее начальному условию

$$(2) \quad Z(x, 0) = 0 \quad (x > 1)$$

и граничному условию

$$(3) \quad Z(1, t) = g(t) \quad (t > 0)$$

(где $g(t)$ — кусочно непрерывно — дифференцируемая функция) с помощью операционного исчисления от Микусински. Сначала они определяют решение $z(x, t)$, относящееся к случаю $g(t) \equiv 1$, в виде

$$z(x, t) = \frac{1}{x} \int_0^t f(\tau) e^{-\frac{x^2}{4(t-\tau)}} d\tau.$$

Эта задача приводит к интегральному уравнению типа свертки относительно неизвестного $f(t)$, не имеющему решения в круге функций. Но это интегральное разрешимо в круге операторов от Mikusiński. Подставив это решение f в приведенное выше выражение для $z(x, t)$, получаем последнее уже в виде функции.

Решение общей проблемы с помощью $z(x, t)$ может быть записано в следующем виде :

$$Z(x, t) = \frac{\partial}{\partial t} \int_0^t g(\tau) z(x, t - \tau) d\tau.$$

В том случае, когда g оператор, не соответствующий функции, решение может быть записано в виде

$$Z(x, t) = s * g * z.$$

PROBLÈMES DE LA PROPAGATION DE LA CHALEUR AVEC DES CONDITIONS AUX LIMITES COMPOSÉES

par

PÁL KOSIK, MELANIE SALLAY et MAGDA ZIMÁNYI

Plaçons une barre ϱ de longueur l homogène conductrice de la chaleur le long de l'axe positif x du système de coordonnées, de telle façon, que l'extrémité $x = 0$ de la barre soit en contact thermique avec un corps β réduit à un point dans lequel la chaleur $Q(t)$ cal/sec se produit en fonction du temps t . Alors en désignant par $V(x, t)$ la distribution de la chaleur de ϱ et par $U(t)$ la température de β , $V(x, t)$ doit satisfaire à l'équation

$$(1) \quad \frac{\partial^2 V}{\partial x^2} = a^2 \frac{\partial V}{\partial t}$$

avec les conditions aux limites et initiales suivantes (voir [2]):

$$(2) \quad AU'(t) + B[U(t) - V(0, t)] = Q(t),$$

$$(3) \quad V_x(0, t) = -C[U(t) - V(0, t)],$$

$$(4) \quad U(0) = U_0, \quad V(x, 0) = f(x),$$

de plus, au cas où $l < \infty$, à l'extrémité $x = l$ de la barre doivent être satisfaites les conditions

$$(a) \quad V_x(l, t) = g(t)$$

ou

$$(b) \quad V(l, t) = h(t),$$

où $Q(t)$, $f(x)$, $g(t)$ et $h(t)$ sont des fonctions données continues. Dans le cas où $l = \infty$ nous supposons que pour un certain γ et pour $t \geq 0$

$$(c) \quad \lim_{x \rightarrow \infty} e^{-\gamma x^2} V(x, t) = 0$$

uniformément en t .

Nous pouvons résoudre le problème de la manière suivante:

Nous déterminons la fonction $V(x, t)$ sous la forme

$$V(x, t) = V^*(x, t) + \bar{V}(x, t),$$

où:

A) $V^*(x, t)$ est la solution de l'équation (1) et satisfait aux conditions (4) et (a) ou (b), de plus $V_x^*(0, t) = 0$. Alors, en définissant les fonctions $U^*(t)$

et $Q^*(t)$ de la manière suivante:

$$U^*(t) = V^*(0, t),$$

$$Q^*(t) = AV_t^*(0, t),$$

et les en substituant dans les expressions (2) et (3), les conditions (2) et (3) sont identiquement satisfaites.

B) $\bar{V}(x, t)$ est la solution de l'équation (1) pour laquelle $\bar{V}(x, 0) = 0$, de plus au cas (a) $\bar{V}_x(l, t) = 0$ au cas (b) $\bar{V}(l, t) = 0$ sont vérifiées et qui satisfait aux conditions aux limites

$$(2) \quad A\bar{U}'(t) + B[\bar{U}(t) - \bar{V}(0, t)] = \bar{Q}(t),$$

$$(3) \quad \bar{V}_x(0, t) = -C[\bar{U}(t) - \bar{V}(0, t)],$$

où

$$\bar{Q}(t) = Q(t) - Q^*(t),$$

$$\bar{U}(t) = U(t) - U^*(t).$$

Dans le cas où $l = \infty$ nous devons exiger que toutes les deux fonctions $V^*(x, t)$ et $\bar{V}(x, t)$ vérifient la condition (c).

G. FREUD dans son travail [2] a démontré que pour toutes les fonctions $\bar{Q}(t)$ continues quelconques la solution du problème B) existe et que nous pouvons obtenir cette solution à l'aide de la transformation de Laplace, en approchant la fonction $Q(t)$ par des polynômes exponentiels de forme

$$\bar{Q}_n(t) = \sum_{k=1}^n a_k e^{-\gamma_k t} \quad (\gamma_k > 0).$$

En effet cherchons la fonction $\bar{V}(x, t)$ sous la forme

$$V(x, t) = \int_0^t F(x, t - \tau) W(\tau) d\tau,$$

où

$$\text{cas (a)} \quad F(x, t) = \frac{1}{\sqrt{t}} \left[e^{-\frac{a^2 x^2}{4t}} + e^{-\frac{a^2 (2l-x)^2}{4t}} \right],$$

$$\text{cas (b)} \quad F(x, t) = \frac{1}{\sqrt{t}} \left[e^{-\frac{a^2 x^2}{4t}} - e^{-\frac{a^2 (2l-x)^2}{4t}} \right],$$

$$\text{cas (c)} \quad F(x, t) = \frac{1}{\sqrt{t}} e^{-\frac{a^2 x^2}{4t}}.$$

$W(t)$ étant choisie arbitrairement la fonction $\bar{V}(x, t)$ satisfait à l'équation (1) et aux conditions aux limites et initiales pour $x = l$ et $t = 0$. Les conditions (2) et (3) fournissent un système d'équations intégrodifférentielles de type

convolutoire pour la détermination des fonctions $U(t)$ et $W(t)$, qui est résoluble à l'aide de la transformation de Laplace au cas où $\bar{U}(0) = \bar{U}_0$, $Q(t) \equiv 0$; $\bar{U}(0) = 0$, $\bar{Q}(t) \equiv 1$ et $\bar{U}(0) = 0$, $\bar{Q}(t) = e^{-\gamma t}$ (concernant la discussion détaillée voir [3]).

Au cas où $\bar{Q}(t) = \sum_{k=1}^{\infty} a_k e^{-\gamma_k t}$, $\bar{U}(0) = U_0$, nous obtenons la solution par superposition des solutions précédentes.

Considérant que $\bar{Q}(t) = Q(t) - Q^*(t)$, la continuité de la fonction $\bar{Q}(t)$ est certainement assurée si $f(x)$ admet une seconde dérivée continue et $f'(0) = 0$.

Dans notre ouvrage nous démontrerons que le problème est aussi résolvable pour $f'(0) = a \neq 0$, c'est à dire pour une $\bar{Q}(t)$ ayant une certaine irrégularité au point $t = 0$, en appliquant la méthode ci-dessus. Avant tout nous démontrerons que dans le cas où $f'(0) = a \neq 0$, $Q(t)$ a une singularité de forme $\frac{\kappa}{\sqrt{t}}$ pour $t = 0$, d'abord nous déterminerons la solution.

Théorème. Soit $V^*(x, t)$ la solution de l'équation de la chaleur sur $0 \leq x \leq l$ pour laquelle $V^*(x, 0) = f(x)$, $V_x^*(0, t) = 0$, de plus au cas où $l \neq \infty$ qui satisfait aux conditions (a) ou (b), ou bien au cas où $l = \infty$ à la condition (c) et soit $f'(0) = a \neq 0$.

Assertion. Dans ce cas on peut écrire $V_t^*(0, t)$ sous la forme $V_t^*(0, t) = \frac{\kappa}{\sqrt{t}} + \psi(t)$, où κ est une constante et $\psi(t)$ est une fonction continue pour $t \geq 0$.

Démonstration. I. Cas d'un intervalle fini.

Soit $r(x, t)$ la solution de l'équation (1) sur l'intervalle $-l \leq x \leq l$, qui vérifie les conditions $r(x, 0) = l - |x|$ et $r(-l, t) = r(l, t) = 0$. En résolvant le problème à l'aide de la méthode de Fourier nous obtenons

$$r(x, t) = \sum_{k=0}^{\infty} \frac{8l}{(2k+1)^2 \pi^2} \cos \frac{(2k+1)\pi x}{2l} e^{-\frac{(2k+1)^2 \pi^2}{(2al)^2} t}.$$

Il en résulte

$$r_t(0, t) = -\frac{2}{a^2 l} \sum_{k=0}^{\infty} e^{-\frac{(2k+1)^2 \pi^2}{(2al)^2} t} = -\frac{2}{a^2 l} \vartheta_2 \left(0 \left| \frac{i\pi}{a^2 l^2} t \right. \right).$$

En appliquant la formule

$$\vartheta_2(z|\tau) = \frac{1}{\sqrt{-i\tau}} e^{-\frac{iz^2}{\tau}} \vartheta_4 \left(-\frac{z}{\tau} \left| -\frac{1}{\tau} \right. \right)$$

nous obtenons

$$r_t(0, t) = -\frac{2}{a^2 l} \vartheta_2 \left(0 \left| \frac{i\pi}{a^2 l^2} t \right. \right) = -\frac{2}{a \sqrt{\pi t}} \vartheta_4 \left(0 \left| i \frac{a^2 l^2}{\tau t} \right. \right).$$

Étant donné que la fonction $\vartheta_4(0|\tau)$ est analytique pour les valeurs $\text{Im } \tau > 0$ (voir [5] p. 476) la fonction $r_t(0, t)$ se produit sous la forme $r_t(0, t) = \frac{\kappa}{\sqrt{t}} A(t)$, où $A(t)$ est analytique. Vu, que le développement infini de la fonction

$\vartheta_4(z|\tau)$ est $\vartheta_4(z|\tau) = 1 - 2q \cos 2z + \dots$, où $q = e^{i\pi\tau}$ nous pouvons écrire la fonction $r_l(0, t)$ sous la forme $r_l(0, t) = \frac{\kappa}{\sqrt{t}} + R(t)$, où $R(t)$ est continue pour $t \geq 0$.

Considérons maintenant la fonction $z(x, t) = V^*(x, t) + a r(x, t)$ sur l'intervalle $0 \leq x \leq l$. La fonction $z(x, t)$ est une solution de l'équation (1) et vérifie les conditions suivantes:

$$z(x, 0) = f(x) + a(l - |x|)$$

$$z_x(0, 0) = 0,$$

au cas (a)

$$z_x(l, t) = g(t) + G(t),$$

où

$$G(t) = 4a \sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{2^{k+1}} e^{-\frac{(2k+1)^2 \pi^2}{(2al)^2} t}$$

et $G(t)$ est continue (voir [4] p. 645), au cas (b)

$$z(l, t) = h(t).$$

La fonction $z(x, t)$ satisfaisant à la condition $z_x(0, 0) = 0$, il découle que $z_l(0, t)$ est continue pour $t \geq 0$. Il en résulte, que $V_l^*(0, t) = z_l(0, t) - a r_l(0, t) = -\frac{a\kappa}{\sqrt{t}} + \psi(t)$, où la fonction $\psi(t) = z_l(0, t) - a R(t)$ est continue pour $t \geq 0$.

II. Cas d'un intervalle infini.

La marche de la démonstration est analogue à la démonstration appliquée dans le cas de l'intervalle fini. Nous devons seulement considérer au lieu de la fonction $r(x, t)$ la fonction $s(x, t)$ pour laquelle $s(x, 0) = e^{-x}$ ($x \geq 0$), $s_x(0, t) = 0$ ($t > 0$) et qui satisfait à l'équation (1) sur $0 \leq x < \infty$.

La solution est

$$s(x, t) = \frac{1}{2} \left\{ e^{\frac{(2t+a^2x)^2 - a^4x^2}{4a^2t}} \left[1 - \Phi \left(\frac{2t + a^2x}{2a\sqrt{t}} \right) \right] + e^{\frac{(2t-a^2x)^2 - a^4x^2}{4a^2t}} \left[1 - \Phi \left(\frac{2t - a^2x}{2a\sqrt{t}} \right) \right] \right\},$$

où

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy.$$

De là

$$s_l(0, t) = -\frac{1}{a\sqrt{\pi t}} + \frac{1}{a^2} e^{\frac{t}{a^2}} \left[1 - \Phi \left(\frac{\sqrt{t}}{a} \right) \right].$$

Il est facile de voir que la fonction $\frac{1}{a^2} e^{\frac{t}{a^2}} \left[1 - \Phi \left(\frac{\sqrt{t}}{a} \right) \right]$ est continue pour $t \geq 0$.

Alors la fonction $s_l(0, t)$ s'écrit sous la forme $-\frac{1}{a\sqrt{\pi t}} + S(t)$, où $S(t)$ est continue.

D'après le théorème il est évident qu'on peut écrire la fonction $\bar{Q}(t) = Q(t) - Q^*(t)$ sous la forme

$$\bar{Q}(t) = H(t) + \frac{c}{\sqrt{t}},$$

où $H(t)$ est continue. Vu, que $\bar{V}(x, t)$ vérifie les conditions homogènes nous obtenons la solution en superposant à la solution appartenant à $\bar{Q}_1(t) = H(t)$ la solution appartenant à $\bar{Q}_2(t) = \frac{c}{\sqrt{t}}$.

La solution du problème étant connue pour $\bar{U}(0) = 0$ et $H(t)$ continue et pour $\bar{U}(0) = U_0$ et $\bar{Q}(t) \equiv 0$, il est suffisant de calculer la fonction $\bar{V}(x, t)$ pour $\bar{U}(0) = 0$ et $\bar{Q}(t) = \frac{1}{\sqrt{t}}$.

Le système des fonctions $\bar{U}(t)$ et $\bar{V}(x, t)$ est univoquement déterminé par les conditions données d'après un corollaire direct du principe de monotonité démontré dans [2]. La démonstration publiée est sans doute applicable dans le cas où $\bar{Q}(t) = ct^a$ ($a > -1$).

En déterminant les fonctions $\bar{U}(t)$, et $\bar{V}(x, t)$ à l'aide de la méthode précédente nous obtenons:

cas (a)

$$\bar{U}(t) = \frac{A_1}{\sqrt{\pi}} \sqrt{t} - 4 \sum_{\varphi_k > 0} r_k E(\varphi_k \sqrt{t}) e^{-\varphi_k^2 t},$$

$$\bar{V}(x, t) = \int_0^t \frac{1}{\sqrt{t-\tau}} \left[e^{-\frac{a^2 x^2}{4(t-\tau)}} + e^{-\frac{a^2 (2l-x)^2}{4(t-\tau)}} \right] W(\tau) d\tau,$$

où

$$W(t) = \frac{A_1}{2} + \sum_{\varphi_k > 0} \left[p_k - \frac{2 s_k}{\sqrt{\pi}} E(\varphi_k \sqrt{t}) \right] e^{-\varphi_k^2 t},$$

cas (b)

$$\bar{U}(t) = 4 \sum_{\varphi_k > 0} r_k E(\varphi_k \sqrt{t}) e^{-\varphi_k^2 t},$$

$$\bar{V}(x, t) = \int_0^t \frac{1}{\sqrt{t-\tau}} \left[e^{-\frac{a^2 x^2}{4(t-\tau)}} - e^{-\frac{a^2 (2l-x)^2}{4(t-\tau)}} \right] W(\tau) d\tau,$$

où

$$W(t) = \sum_{\varphi_k > 0} \left[p_k - \frac{2 s_k}{\sqrt{\pi}} E(\varphi_k \sqrt{t}) \right] e^{-\varphi_k^2 t},$$

de plus

$$\begin{aligned} r_k &= \frac{\beta_1}{\alpha_1 \varphi_k^5 + \alpha_2 \varphi_k^3 + \alpha_3 \varphi_k} & A_1 &= \frac{C}{AC + Ba^2 l} \\ p_k &= \frac{\beta_2 \varphi_k^2 + \beta_3}{\alpha_1 \varphi_k^4 + \alpha_2 \varphi_k^2 + \alpha_3} & \alpha_1 &= -A^2 a^2 l \\ s_k &= \frac{\beta_4}{\alpha_1 \varphi_k^4 + \alpha_2 \varphi_k^2 + \alpha_3} & \alpha_2 &= A[-AC^2 l - AC + 2Ba^2 l] \\ \beta_1 &= BC & \beta_3 &= -BCa & \alpha_3 &= -B(AC + Ba^2 l) \\ \beta_2 &= ACa & \beta_4 &= AC^2 & E(x) &= \int_0^x e^{y^2} dy \end{aligned}$$

Les valeurs φ_k sont les racines des équations suivantes:
au cas (a)

$$\operatorname{tg} al \varphi = \frac{AC \varphi}{a(A \varphi^2 - B)},$$

au cas (b)

$$\operatorname{tg} al \varphi = \frac{a(B - A \varphi^2)}{AC \varphi}.$$

cas (c) ($l = \infty$)

$$\overline{U}(t) = \frac{c\sqrt{\pi}}{Ba} + 2\gamma_1 e^{p_1^2 t} \operatorname{Erf}(-p_1 \sqrt{t}) - 2\gamma_2 e^{p_2^2 t} \operatorname{Erf}(-p_2 \sqrt{t}),$$

$$\overline{V}(x, t) = \int_0^t \frac{1}{\sqrt{t-\tau}} e^{-\frac{a^2 x^2}{4(t-\tau)}} W(\tau) d\tau,$$

où

$$W(t) = \frac{2}{\sqrt{\pi}} \frac{C}{p_1 - p_2} [e^{p_1^2 t} \operatorname{Erf}(-p_1 \sqrt{t}) - e^{p_2^2 t} \operatorname{Erf}(-p_2 \sqrt{t})]$$

de plus

$$\gamma_1 = \frac{ap_1 + C}{p_1(p_2 - p_1)}$$

$$\gamma_2 = \frac{ap_2 + C}{p_2(p_2 - p_1)}$$

$$\operatorname{Erf}(x) = \int_x^\infty e^{-y^2} dy.$$

Ici p_1 et p_2 sont les racines supposées différentes de l'équation $Aap^2 + ACp + Ba = 0$.
Considérant que

I. la fonction $V(x, t)$ admet une seconde dérivée continue pour $t > 0$ et satisfait à l'équation (1), de plus aux conditions prescrites pour $t = 0$ et $x = l$;

II. $W(t)$ est continue et $\lim_{x \rightarrow 0} \frac{\partial}{\partial x} \bar{V}(x, t)$ existe pour $t > 0$ (voir [1]);

III. $\bar{U}(t)$ est continue et dérivable et $\bar{U}'(t)$ prend la forme $\frac{\text{const}}{\sqrt{t}} + K(t)$ (où $K(t)$ est continue),

les fonctions obtenues sont les solutions du problème.

Enfin nous exprimons nos vifs remerciements à M. G. FREUD pour l'aide précieuse apportée dans notre ouvrage.

(Reçu le 9 octobre 1959.)

BIBLIOGRAPHIE

- [1] ADLER Gy.: „Hővezetési és diffúziós feladatok összetett peremfeltételekkel”, II. *A Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei* **1** (1956) 167—183.
- [2] FREUD, G.: „Problèmes de la propagation de la chaleur avec les conditions aux limites composées”. (Sous la presse.)
- [3] FREUD G.: „Hővezetési és diffúziós feladatok összetett peremfeltételekkel, I.” *A Magyar Tudományos Akadémia Alkalmazott Matematikai Intézetének Közleményei* **3** (1955) 369—394.
- [4] SZÁSZ P.: *A differenciál- és integrálszámítás elemei. I. 2. kiadás.* Közoktatásügy Kiadóvállalat, Budapest, 1951.
- [5] WHITTAKER, E. T.—WATSON, G. N.: *A course of modern analysis.* 4. edition. University Press, Cambridge, 1952.

HŐVEZETÉSI PROBLÉMÁK ÖSSZETETT HATÁRFELTÉTELEK ESETÉN

KOSIK P., SALLAY M. és ZIMÁNYI M.

Kivonat

A szerzők az (1) egyenlet megoldásával foglalkoznak, a (2), (3) és (4), valamint az (a), (b) vagy (c) feltételek fennállása mellett. Freud G. [2] dolgozatában megadja a feladat megoldását az $f'(0) = 0$ feltétel mellett. A szerzők jelen dolgozatukban megadják a megoldást ezen korlátozó feltevés nélkül.

ПРОБЛЕМЫ ТЕПЛОПРОВОДНОСТИ В СЛУЧАЕ СЛОЖНЫХ ГРАНИЧНЫХ УСЛОВИЙ

P. KOSIK, M. SALLAY и M. ZIMÁNYI

Резюме

Авторы занимаются решением уравнения (1) в случае выполнения условий (2), (3) и (4), а также (a), (b) или (c). G. FREUD в работе [2] решил задачу при условии $f'(0) = 0$. Авторы в настоящей работе дают решение без этого ограничивающего условия.

ON CONNECTED GRAPHS, I.

by

A. RÉNYI

Let $G_{n,N}$ denote a connected graph having the n labelled vertices P_1, P_2, \dots, P_n and N edges. Let $C(n, N)$ denote the number of all possible $G_{n,N}$. Let us call the number $d = N - n + 1$ the „degree of connectivity” of the connected graph $G_{n,N}$. Clearly $d \geq 0$, and the graph is a tree if and only if $d = 0$. The number of different trees with n labelled vertices is, according to a classical result of CAYLEY [1] equal to n^{n-2} , i. e.

$$(1) \quad C(n, n-1) = n^{n-2} \quad (n = 1, 2, \dots).$$

A simple explicit formula for $C(n, N)$ is not known, moreover the asymptotic behaviour of $C(n, n+d-1)$, if d is fixed and $n \rightarrow +\infty$ has according to the knowledge of the author, not determined up to now.

One can give some rather complicated explicit formulae resp. recursive relations for $C(n, N)$ (see [2] and [3]). For instance one obtains by the usual sieve-method

$$(2) \quad C(n, N) = \binom{n}{2} - \frac{1}{2} \sum_{k=1}^{n-1} \binom{n}{k} \left(\binom{k}{2} + \binom{n-k}{2} \right) + \\ + \frac{1}{3} \sum_{k+l < n} \frac{n!}{k! l! (n-k-l)!} \left(\binom{k}{2} + \binom{l}{2} + \binom{n-k-l}{2} \right) - \dots$$

which leads to an explicit formula for the generating function

$$(3) \quad \sum_{n=1}^{\infty} \sum_{N=1}^{\infty} \frac{c(n, N) x^n y^N}{n!} = \log \left(1 + \sum_{k=1}^{\infty} \frac{(1+y)^{\binom{k}{2}} x^k}{k!} \right).$$

There is also the recursion formula

$$(4) \quad \binom{n+1}{2} = \sum_{k=0}^n \binom{n}{k} \sum_{m=k}^{\binom{k+1}{2}} \binom{n-k}{2} C(k+1, m).$$

It seems however that these formulae do not help much if one wants to determine the asymptotic behaviour of $C(n, n+d-1)$.

The aim of the present note is to discuss the case $d = 1$ that is to give for $C(n, n)$ an asymptotic formula. In a subsequent note we shall discuss the cases $d \geq 2$. Clearly if a connected graph consists of n vertices and n edges, it contains exactly one circle, some vertices of which are the roots of a tree, these trees being disjoint. (Fig. 1. shows all possible types of connected graphs with 6 vertices and 6 edges.)

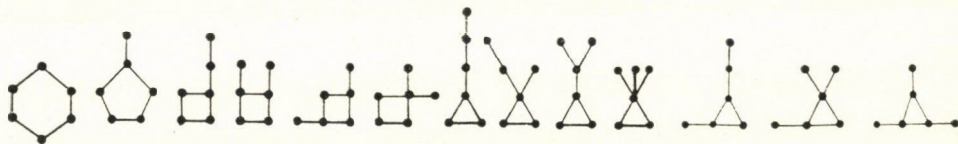


Figure 1.

If the simple circle contained in such a graph consists of k vertices, then $3 \leq k \leq n$ all these values being possible. If the edges of this circle are removed from a graph of the considered type there remains a graph with n vertices consisting of k disjoint trees as subgraphs such that the k vertices of the removed circle belong to different subgraphs. The number $T(n, k)$ of such graphs has been determined already by CAYLEY [3]; he asserted that

$$(4) \quad T(n, k) = kn^{n-k-1}.$$

A proof of this result of Cayley has been given in [4]. This formula is contained as a special case in a more general result of G. W. FORD and G. E. UHLENBECK [5] (see also [6]). Using this result one easily obtains

$$(5) \quad C(n, n) = \sum_{k=3}^n T(n, k) \binom{n}{k} \frac{(k-1)!}{2} = \frac{n^{n-1}}{2} \sum_{k=3}^n \binom{n}{k} \frac{k!}{n^k}.$$

As a matter of fact the vertices of a circle consisting of k points can be chosen in $\binom{n}{k}$ ways; they can be arranged to a circle in $\frac{(k-1)!}{2}$ different ways; the remaining points can be arranged as mentioned above in $T(n, k)$ different ways; as k may take any value from 3 to n , formula (5) follows immediately from (4). (5) can be written also in the form

$$(5') \quad C(n, n) = \frac{1}{2} n^{n-1} \sum_{k=3}^n \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right).$$

Using (5') it is easy to determine the asymptotic behaviour of $C(n, n)$. As a matter of fact

$$\prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right) \leq e^{-\frac{1}{n} \binom{k}{2}}$$

and for $k = o(n^{2/3})$

$$(6) \quad \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right) = e^{-\frac{k^2}{2n}} \left(1 + O\left(\frac{k^3}{n^2}\right)\right).$$

Thus it follows that

$$(7) \quad \sum_{k=3}^n \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right) = \sqrt{\frac{n\pi}{2}} + O(1)$$

and therefore

$$(8) \quad C(n, n) \sim \sqrt{\frac{\pi}{8}} n^{n-\frac{1}{2}}.$$

It is interesting to determine the distribution of the length of the circle contained in a random $G_{n,n}$.

If γ_n denotes the length of the circle contained in a $G_{n,n}$ chosen at random (so that each of the $C(n, n)$ graphs $G_{n,n}$ has the same probability to be chosen) we obtain

$$(9) \quad \mathbf{P}\left(\frac{\gamma_n}{\sqrt{n}} < x\right) = \frac{1}{2} n^{n-1} \sum_{3 \leq k < x\sqrt{n}} \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right).$$

It follows by some elementary calculation, using (6), that

$$(10) \quad \lim_{n \rightarrow +\infty} \mathbf{P}\left(\frac{\gamma_n}{\sqrt{n}} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{+x} e^{-\frac{u^2}{2}} du.$$

Thus $\frac{\gamma_n}{\sqrt{n}}$ has in the limit for $n \rightarrow +\infty$ the same distribution as the absolute value of a random variable having a normal distribution with mean 0 and variance 1. It follows that the mean value of γ_n is asymptotically $\sqrt{\frac{2n}{\pi}}$.

The results of the present note will be used in a forthcoming joint paper of P. ERDŐS and the author.

(Received October 26, 1959.)

REFERENCES

- [1] CAYLEY, A.: „A theorem on trees.” *Quarterly Journal of Pure and Applied Mathematics* **23**(1889) 376—378.
- [2] RIDDELL, R. I.—UHLENBECK, JR. G. E.: „On the theory of the virial development of the equation of state of monoatomic gases.” *Journal of Chemical Physics* **21** (1953) 2056—2064.
- [3] GILBERT, E. N.: „Enumeration of labelled graphs.” *Canadian Journal of Mathematics* **8** (1956) 405—411.
- [4] RÉNYI, A.: „Some remarks on the theory of trees.” *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* **4**(1959) 73—85.
- [5] FORD, G. W.—UHLENBECK, G. E.: „Combinatorial problems in the theory of graphs, I.” *Proceedings of the National Academy of Sciences of the USA*. **42** (1956) 122—128.
- [6] AUSTIN, T. L.—FAGEN, R. E.—PENNEY, W. F.—RIORDAN, J.: „The number of components in random linear graphs.” *Annals of Math. Statistics* **30**(1959) 747—754.

Remark added on March 21, 1960.

The following reference should be added: Formula (5) is contained (with an other interpretation) in the paper by L. KATZ: "Probability of indecomposability of a random mapping function" (Annals of Mathematical Statistics 26 (1955) 512—517), where the number $I(n)$ of indecomposable single-valued mappings of a set of n (labelled) elements into itself is determined. It is easy to see that $I(n) = 2C(n, n)$, which establishes the equivalence of the theorem of KATZ with formula (5). The asymptotic formula (8) is also contained in the mentioned paper of KATZ.

ÖSSZEFÜGGŐ GRÁFOKRÓL, I.

RÉNYI A.

Jelölje $C(n, N)$ az n (számozott) pontból és N élből álló összes lehetséges összefüggő gráfok számát. Felhasználva CAYLEY egy képletét (amelynek bizonyítása [4]-ben található meg) a szerző kiszámítja $C(n, n)$ -et (azaz azon összefüggő n -csúcspontú gráfok számát, amelyek egyetlen kört tartalmaznak)

és kimutatja, hogy $C(n, n) \sim \sqrt{\frac{\pi}{8}} n^{n-1/2}$. E dolgozat folytatásában a szerző $C(n, n+d)$ aszimptotikus viselkedését fogja vizsgálni rögzített d mellett $n \rightarrow +\infty$ esetében.

О СВЯЗНЫХ ГРАФАХ, I.

A. RÉNYI

Резюме

Пусть $C(n, N)$ означает число всех возможных связных графов с n нумерованными точками и N ребрами. Применяя одну формулу от CAYLEY (доказанную в [4]) автор вычисляет $C(n, n)$ и покажет что $C(n, n) \sim \sqrt{\frac{\pi}{8}} n^{n-1/2}$.

В продолжении настоящей заметки автор будет заниматься с определением асимптотической поведении от $C(n, n+d)$ для фиксированной $d \geq 1$ для $n \rightarrow +\infty$.

SUMMATION METHODS AND PROBABILITY THEORY

by

A. RÉNYI

§ 1. Probabilistic interpretation of methods of summation

Let $\mathbf{A} = (a_{nk})$ be an infinite matrix with *nonnegative* elements, with row-sums equal to 1 and such that the elements of each column tend to 0, i. e.

$$(1.1) \quad a_{nk} \geq 0 \quad (n, k = 0, 1, 2, \dots)$$

$$(1.2) \quad \sum_{k=0}^{\infty} a_{nk} = 1 \quad (n = 0, 1, \dots)$$

and

$$(1.3) \quad \lim_{n \rightarrow +\infty} a_{nk} = 0 \quad (k = 0, 1, \dots).$$

As well known (see [1]), the summation method which consists in forming from a given sequence $s_k (k = 0, 1, \dots)$ the transformed sequence $t_n = \mathbf{A} s_n$ defined by

$$(1.4) \quad t_n = \sum_{k=0}^{\infty} a_{nk} s_k \quad (n = 0, 1, \dots)$$

and considering the limit of t_n (if it exists), is *permanent*, i. e. if $\lim_{n \rightarrow +\infty} s_n = s$ then $\lim_{n \rightarrow +\infty} t_n = s$ too. Such a method can be interpreted probabilistically as follows: let $v_n (n = 0, 1, \dots)$ be a sequence of random variables, taking on only nonnegative integral values, with the corresponding probabilities

$$(1.5) \quad \mathbf{P}(v_n = k) = a_{nk} \quad (n, k = 0, 1, \dots).$$

(Here and in what follows we denote by $\mathbf{P}(\dots)$ the probability of the event in the brackets.) The conditions (1.1) and (1.2) clearly express only that the sequence $a_{nk} (k = 0, 1, \dots)$ is for each value of n the probability distribution of such a random variable. Condition (1.3) expresses that v_n tends in probability to $+\infty$ (which we denote by $v_n \Rightarrow +\infty$); as a matter of fact by (1.5) and (1.3) we have

$$(1.6) \quad \lim_{n \rightarrow +\infty} \mathbf{P}(v_n \geq N) = 1$$

for all positive values of N , which is equivalent with $v_n \Rightarrow +\infty$. Now (1.4) can be interpreted as follows: t_n is the mean value of the random variable s_{v_n} that is

$$(1.7) \quad t_n = \mathbf{M}(s_{v_n}).$$

(Here and in what follows $\mathbf{M}(\dots)$ denotes the mean value of the random variable in the brackets.) Thus the summation procedure defined by the matrix $\mathbf{A} = (a_{nk})$ satisfying (1.1) — (1.2) — (1.3), can be interpreted as follows: we consider the mean value of the random term s_{v_n} of the sequence s_k and consider the limit of this mean value if $n \rightarrow +\infty$.

§ 2. Hausdorff-methods

For HAUSDORFF-summation methods, still more can be said. The summation corresponding to the matrix \mathbf{A} is called a HAUSDORFF-method of summation if

$$(2.1) \quad a_{nk} = \binom{n}{k} \int_0^1 x^k (1-x)^{n-k} dF(x)$$

where $F(x)$ is a probability distribution function in the interval $[0, 1]$; the process is as well known (see [1]) permanent if $F(x)$ is continuous in the point $x = 0$.

In case of a *Hausdorff*-method the random variables v_n introduced by (1.5) can be characterized as follows: $v_n = \beta(n, \xi)$, where ξ is a random variable which has the distribution function $F(x)$, and $\beta(n, x)$ is a random variable for each fixed value of x ($0 < x < 1$) which has a binomial distribution of order n with parameter x , i. e.

$$(2.2) \quad \mathbf{P}(\beta(n, x) = k) = \binom{n}{k} x^k (1-x)^{n-k} \quad (k = 0, 1, \dots, n).$$

Thus, we first choose a value of the random variable ξ ; if this value is x we draw with replacement n balls from an urn containing red and white balls in the proportion x to $1-x$, and if the number of red balls among the balls chosen is k , we put $v_n = k$. Thus we have

$$(2.3) \quad t_n = \mathbf{M}(s_{\beta(n, \xi)}).$$

This interpretation enables us to prove a number of known facts about HAUSDORFF-summation methods in a surprisingly simple manner. For instance let us consider two HAUSDORFF-methods corresponding to the matrices $\mathbf{A} = (a_{nk})$ and $\mathbf{B} = (b_{nk})$ where a_{nk} is defined by (2.1) and b_{nk} by

$$(2.4) \quad b_{nk} = \binom{n}{k} \int_0^1 x^k (1-x)^{n-k} dG(x)$$

where $G(x)$ is an other distribution function in $[0, 1]$. Let us now perform the two transformations $t_n = \mathbf{A} s_n$ and $v_n = \mathbf{B} t_n$ after another, i. e. consider

$v_n = \mathbf{BA} s_n$. It is well known that the resulting method is again one of the Hausdorff type. This can be shown as follows: If η is a random variable having $G(x)$ for its distribution function, and independent of ξ , we have by (2.4)

$$(2.5) \quad v_n = \mathbf{M}(\mathbf{M}(s_{\beta(\beta(n, \eta), \xi)})) .$$

Now the double mean value can be replaced according to a known theorem of probability theory (see [2]) by a simple one, i. e.

$$(2.6) \quad v_n = \mathbf{M}(s_{\beta(\beta(n, \eta), \xi)}) .$$

Thus we have to proceed as follows: first observe the values of the random variables ξ and η which are independent and have the distribution functions $F(x)$ and $G(x)$; if these values are $\xi = x$ and $\eta = y$, then take at random (with replacement) n balls from an urn which contains red and white balls in the proportion y to $1 - y$. If among the balls there are k red ones, take at random (with replacement) k balls from an urn composed of red and white balls in the proportion x to $1 - x$; if among the balls chosen there are l red ones, put $v_n = l$. Now by a well known theorem of probability, the distribution of v_n is the same as if we would have used only one urn which is composed of red and white balls in the proportion xy to $1 - xy$ and choose at random (with replacement) n balls.

This fact can be expressed by saying that the mixture of binomial distributions with the same parameter x and of different orders, with weights forming also a binomial distribution with parameter y and order n , is itself a binomial distribution of order n and parameter xy . This can be proved e. g. by the symbolical calculus of distributions (see [3], p. 129—133) as follows:

$$(2.7) \quad \sum_{k=0}^n \binom{n}{k} y^k (1-y)^{n-k} (xE_1 + (1-x))^k = (xy E_1 + (1-xy))^n$$

where E_1 denotes the distribution attributing the probability 1 to the value 1. Thus it follows that

$$(2.8) \quad v_n = \mathbf{M}(s_{\beta(n, \xi \eta)}) .$$

Thus $v_n = \mathbf{C} s_n$ where \mathbf{C} is the HAUSDORFF-matrix with the elements

$$(2.9) \quad c_{nk} = \binom{n}{k} \int_0^1 x^k (1-x)^{n-k} dH(x)$$

where $H(x)$ is the distribution function of the random variable $\xi \eta$, that is

$$(2.10) \quad H(x) = \int_0^1 F\left(\frac{x}{u}\right) dG(u) = \int_0^1 G\left(\frac{x}{u}\right) dF(u) .$$

Clearly the continuity of either $F(x)$ or $G(x)$ in $x = 0$ implies the same for $H(x)$.

§ 3. Henriksson-methods

If we replace in the definition of the HAUSDORFF-method the binomial distribution $\left\{ \binom{n}{k} x^k (1-x)^{n-k} \right\}$ by the Poisson-distribution $\left\{ \frac{(\lambda x)^k e^{-\lambda x}}{k!} \right\}$ where $\lambda > 0$, we obtain an other important class of summation methods: *Henriksson's class of methods of summation*. The definition of this class of summation methods given below is slightly different from that given by HENRIKSSON (see [5]).

Let $F(x)$ be a distribution function in the interval $[0, +\infty)$; there corresponds to every $F(x)$ a method of summation which is defined as follows: we put

$$(3.1) \quad a_k(\lambda) = \int_0^{\infty} \frac{(\lambda x)^k}{k!} e^{-\lambda x} dF(x) \quad (k = 0, 1, \dots)$$

form the transformed values

$$(3.2) \quad t(\lambda) = \sum_{k=0}^{\infty} a_k(\lambda) s_k$$

and consider the limit

$$(3.3) \quad \lim_{\lambda \rightarrow +\infty} t(\lambda) = s.$$

If the limit (3.3) exists we shall say that the sequence $\{s_n\}$ is summable to s by the *Henriksson-method of summation corresponding to the distribution function $F(x)$* . If $\gamma(\lambda)$ denotes a random variable having Poisson-distribution with mean value λ , then clearly

$$(3.4) \quad t(\lambda) = \mathbf{M}(s_{\gamma(\lambda)\xi})$$

where ξ is a random variable having the distribution function $F(x)$. It is easy to see that the Henriksson method corresponding to the distribution function $F(x)$ is permanent if and only if $F(x)$ is continuous for $x = 0$.

The class of Henriksson methods includes of course Borel's methods, which is obtained if $F(x)$ is the distribution function of the constant 1 in which case

$$(3.5) \quad t(\lambda) = \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} s_k.$$

The class of Henriksson's summation methods includes further Abel's method, which is obtained if $F(x)$ is an exponential distribution function, $F(x) = 1 - e^{-x}$

for $x \geq 0$ in which case, putting $u = \frac{\lambda}{1 + \lambda}$ we have

$$(3.6) \quad t(\lambda) = (1 - u) \sum_{k=0}^{\infty} s_k u^k.$$

If $\lambda \rightarrow +\infty$ then clearly $u \rightarrow 1 - 0$, thus $t(\lambda) \rightarrow s$ means nothing else than the Abel-summability of s_n to s .

If we choose for $F(x)$ the gamma-distribution of order r

$$(3.7) \quad F(x) = \int_0^x \frac{t^{r-1} e^{-t}}{(r-1)!} dt \quad (r > 0)$$

we obtain putting again $U = \frac{\lambda}{1 + \lambda}$

$$(3.8) \quad t(\lambda) = (1 - u)^r \sum_{k=0}^{\infty} \binom{k+r-1}{r-1} u^k s_k$$

which is a generalization of the Abel-method and may be called the Abel-method of order r ($r = 1$ corresponds to the ordinary Abel-method of summation).

O. HENRIKSSON defined his class of summation methods in a different way, by putting

$$(3.9) \quad t(\lambda) = \sum_{k=0}^{\infty} s_k \lambda^k \frac{q^{(k)}(-\lambda)}{k!}$$

where he supposed that $q(z)$ is a power series

$$(3.10) \quad q(z) = \sum_{n=0}^{\infty} q_n z^n.$$

Clearly, putting

$$(3.11) \quad q(-z) = \int_0^{\infty} e^{-zx} dF(x) \quad \text{for } z \geq 0$$

we have

$$(3.12) \quad \frac{\lambda^k q^{(k)}(-\lambda)}{k!} = a_k(\lambda).$$

Thus the of summations-methods defined by (3.9) is essentially the same as the class defined by (3.2). Our treatment is somewhat more general then that of HENRIKSSON, as we do not suppose that $q(z)$ is regular in the point $z = 0$, only that it is of the form (3.11) where $F(x)$ is a distribution function, i. e. that $q(-z)$ is the Laplace-transform of a distribution function in the interval $(0, +\infty)$. By other words we suppose that $q(-z)$ is completely monotonic for $z \geq 0$ and $q(0) = 1$ as by a well known theorem of HAUSDORFF (see e. g.

[6], p. 89, Theorem 3.5.) it follows that $q(-z) = \int_0^{\infty} e^{-zx} dF(x)$ where $F(x)$ is a distribution function, further that $\lim_{z \rightarrow +\infty} q(-z) = 0$, which implies that $F(x)$ is continuous at $x = 0$. On the other hand Henriksson did not suppose that the coefficients q_n in (3.10) are non-negative.

It has been already remarked by HENRIKSSON (see [5]) that if one performs on $\{s_n\}$ a Hausdorff-transformation and after this a Henriksson-transformation, this is equivalent to a single Henriksson-transformation. As a matter of fact if ξ has the distribution function $F(x)$ ($0 \leq x \leq 1$) and η the

distribution function $G(y)$ ($0 \leq y \leq +\infty$) and ξ and η are independent then putting $t_n = \mathbf{M}(s_{\beta(n, \xi)})$ and $v(\lambda) = \mathbf{M}(t_{\gamma(\lambda, \eta)})$ one has

$$(3.13) \quad v(\lambda) = \mathbf{M}(\mathbf{M}(s_{\beta(\gamma(\lambda, \eta), \xi)})) = \mathbf{M}(s_{\beta(\gamma(\lambda, \eta), \xi)}).$$

Now as well known from probability theory, the mixture of binomial distributions by Poissonian weights is again a Poisson distribution. This can be shown by the symbolical calculus as follows:

$$(3.14) \quad \sum_{n=0}^{\infty} \frac{\mu^n}{n!} e^{-\mu} (p E_1 + (1-p))^n = e^{\mu p(E_1 - 1)}.$$

Thus we have

$$(3.15) \quad \mathbf{M}(s_{\beta(\gamma(\lambda, \eta), \xi)}) = \mathbf{M}(s_{\gamma(\lambda, \xi \eta)}).$$

Thus putting

$$(3.16) \quad b_k(\lambda) = \int_0^{\infty} \frac{(\lambda x)^k}{k!} e^{-\lambda x} dH(x)$$

where

$$(3.17) \quad H(z) = \int_0^1 G\left(\frac{z}{x}\right) dF(x) = \int_0^{\infty} F\left(\frac{z}{x}\right) dG(x)$$

we have

$$(3.18) \quad v(\lambda) = \sum_{k=0}^{\infty} b_k(\lambda) s_k.$$

Thus $v(\lambda)$ is obtained by using the Henriksson-transformation corresponding to the distribution function $H(x)$ given by (3.17).

§ 4. Limit-distribution methods of summation

It is a natural idea to try to characterize the „sum” of a divergent series by a distribution instead of by a number. This can be done in different ways. For instance if $\{s_k\}$ is a sequence of real numbers, let us put

$$(4.1) \quad S(y) = \lim_{n \rightarrow +\infty} \frac{1}{n+1} \sum_{\substack{k \leq n \\ k < y}} 1$$

provided that the limit on the right of (4.1) exists. If $S(y)$ is a distribution function and convergence in (4.1) takes place for all points of continuity y of $S(y)$, we call $S(y)$ the $(C, 1)$ -limiting distribution of the sequence $\{s_n\}$. The relation (4.1) can be put also in the following form:

$$(4.2) \quad S(y) = \lim_{n \rightarrow +\infty} \frac{1}{n+1} \sum_{k=0}^n H(y - s_k)$$

where $H(x)$ is the Heaviside function

$$H(x) = \begin{cases} 1 & \text{for } x > 0 \\ 0 & \text{for } x \leq 0. \end{cases}$$

Clearly if (4.1) exists for any real y , and $f(x)$ is an arbitrary step-function, we have

$$(4.3) \quad \lim_{n \rightarrow +\infty} \frac{1}{n+1} \sum_{k=0}^n f(s_k) = \int_{-\infty}^{+\infty} f(y) dS(y).$$

Formula (4.3) shows the real meaning of the limit distribution function $S(y)$, further shows the way of its natural generalization. If A is an arbitrary linear method of summation, we may consider the limit (if it exists)

$$(4.4) \quad S(y) = A - \lim_{n \rightarrow +\infty} H(y - s_n)$$

and if $S(y)$ is a probability distribution function, call it the A -limiting distribution function of the sequence $\{s_n\}$. (4.4) implies that for any step function f we have

$$(4.5) \quad A - \lim_{n \rightarrow +\infty} f(s_n) = \int_{-\infty}^{+\infty} f(y) dS(y).$$

Of course (4.3) holds also for any Riemann-integrable function $f(x)$ for which $\int_{-\infty}^{+\infty} f(x) dS(x)$ exists, as for such a function and for any $\varepsilon > 0$ two step functions $f_1(x)$ and $f_2(x)$ can be found so that

$$f_1(x) \leq f(x) \leq f_2(x) \quad \text{and} \quad \int_{-\infty}^{+\infty} (f_2(x) - f_1(x)) dS(x) < \varepsilon.$$

Concerning (4.5) the same holds if A is a linear method having a matrix with nonnegative elements.

Conversely, if the limit in (4.3) resp. (4.5) exists for $f(x) = e^{ixt}$ for all real t and if it is a continuous function of t for $t = 0$, then by the well known theorem on characteristic functions (4.2) resp. (4.4) holds.

If (4.2) holds and $\int_{-\infty}^{+\infty} x dS(x)$ is finite, we may consider $s = \int_{-\infty}^{+\infty} x dS(x)$ as the generalized limit of the sequence $\{s_n\}$. If moreover s_n is bounded, then clearly $\lim_{n \rightarrow +\infty} \frac{1}{n+1} \sum_{k=0}^n s_k = \int_{-\infty}^{+\infty} x dS(x)$. If s_n is unbounded, this is evidently not true, as $S(y)$ remains unchanged if s_n is arbitrarily changed for a sequence of values of n which have 0 density, while by such an operation the $(C, 1)$ -summability can be clearly destroyed.

Note that $S(y)$ is the distribution function of a constant s , i. e.

$$(4.6) \quad S(y) = \begin{cases} 0 & \text{for } y \leq s \\ 1 & \text{for } y > s \end{cases}$$

if and only if s_n is almost convergent (see [7], Vol. II, p. 181), i. e. the sequence of positive integers can be split into two disjoint sequences m_k and l_k ($k = 1, 2, \dots$) such that $\lim_{k \rightarrow +\infty} s_{m_k} = s$ and the sequence l_k has zero density, i. e.

$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{l_k < N} 1 = 0$. Thus the existence of a limiting distribution of a sequence may be considered as the generalization of almost convergence, not of the ordinary convergence.

§ 5. On the sequence of generalized partial sums of a sequence

In a recent paper [3] I considered the following problem. Let $\sum_{k=0}^{\infty} a_k$ be a series, and denote by A_n the sum

$$(5.1) \quad A_n = a_{k_1} + a_{k_2} + \dots + a_{k_r}$$

provided that

$$(5.2) \quad n = 2^{k_1} + 2^{k_2} + \dots + 2^{k_r} \quad (k_1 > k_2 > \dots > k_r)$$

is the representation of n in the binary number system.

The sequence $\{A_n\}$ contains clearly all finite sums formed from the terms of the series $\sum a_k$, arranged in lexicographic order. It has been shown in [3] that the sequence $\{A_n\}$ has a $(C, 1)$ -limiting distribution in the sense of § 4. if and only if both $\sum a_k$ and $\sum a_k^2$ are convergent. It has been proved in [3] also that the sequence $\{A_n\}$ is $(C, 2)$ -summable if and only if the series $\sum a_k$ is convergent.

I conjectured that the same holds for Abel-summability, i. e. we have

$$(5.3) \quad \lim_{x \rightarrow 1-0} (1-x) \sum_{n=0}^{\infty} A_n x^n = s$$

if and only if the series $\sum a_k$ is convergent and

$$(5.4) \quad \sum_{k=0}^{\infty} a_k = 2s.$$

That (5.4) implies (5.3) is easily shown. However I did not succeed up to now to prove that (5.3) implies (5.4). This would follow from the following equivalent statement:

If

$$(5.5) \quad \lim_{x \rightarrow 1-0} \sum_{k=0}^{\infty} a_k \frac{x^{2^k}}{1+x^{2^k}} = s$$

then the series $\sum_{k=0}^{\infty} a_k$ is convergent and its sum is $2s$.

This „high-indices”-type conjecture (5.5) is up to now not proved. It is easy to prove however (5.5) if some additional Tauberian condition is supposed, e. g. if we suppose

$$(5.6) \quad \lim_{k \rightarrow +\infty} a_k = 0.$$

As a matter of fact, if (5.6) holds then (5.5) follows from the „high-indices” theorem (see [1], Theorem 114.).

As a matter of fact, if (5.5) and (5.6) hold, we have

$$\left| \sum_{k=0}^{\infty} \frac{2a_k x^{2^k}}{1+x^{2^k}} - \sum_{k=0}^{\infty} a_k x^{2^k} \right| \leq \sum_{k=0}^{\infty} |a_k| x^{2^k} (1-x^{2^k})$$

and thus, if $\text{Max}_{k \geq n} |a_k| = B_n$, then

$$(5.7) \quad \left| \sum_{k=0}^{\infty} \frac{2 a_k x^{2^k}}{1 + x^{2^k}} - \sum_{k=0}^{\infty} a_k x^{2^k} \right| \leq B_0(1 - x^{2^n}) + B_n.$$

Now let $x \rightarrow 1 - 0$ and $n \rightarrow +\infty$ but so that $x^{2^n} \rightarrow 1$. Then the right-hand side of (5.7) tends to 0. Thus it follows from (5.5) that

$$(5.8) \quad \lim_{x \rightarrow 1-0} \sum_{k=0}^{\infty} a_k x^{2^k} = 2s$$

and by the high-indices theorem it follows that $\sum_{k=0}^{\infty} a_k$ is convergent with the sum $2s$.

(Received October 26, 1959.)

REFERENCES

- [1] HARDY, G. H.: *Divergent series*. Oxford, 1949.
- [2] KOLMOGOROFF, A.: *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin, 1933.
- [3] RÉNYI A.: *Valószínűségszámítás*. (Textbook of probability theory), Tankönyvkiadó, Budapest, 1954.
- [4] RÉNYI, A.: „Mathematical Notes, II. On the sequence of generalized partial sums of a series”. *Publicationes Math.* **5** (1957) 129—141.
- [5] HENRIKSSON, O.: „Über die Hausdorffsche Limitierungsverfahren die schwächer sind als das Abelsche”. *Math. Zeitschr.* **39** (1935) 501—510.
- [6] SHOHAT, I. A.—TAMARKIN, J. D.: „The problem of moments”. *Math. Surveys*, No. 1. American Math. Soc. 1950.
- [7] ZYGMUND, A.: *Trigonometric series*. 2nd. ed. Cambridge University Press, 1959

Remark added on March 21, 1960.

The following reference should be mentioned: D. D. KOSAMBI in his paper “Classical Tauberian theorems” (*Journal of the Indian Society of Agricultural Statistics* 10 (1958) 141—149) makes use of the same probabilistic interpretation of methods of summation as given in § 1.

SZUMMÁCIÓS ELJÁRÁSOK VALÓSZÍNŰSÉGSZÁMÍTÁSI INTERPRETÁCIÓJA

RÉNYI A.

Legyen $\mathbf{A} = (a_{nk})$ ($n, k = 0, 1, \dots$) egy végtelen mátrix, amelyre teljesülnek a következő feltételek:

$$(1.1) \quad a_{nk} \geq 0 \quad (n, k = 0, 1, \dots),$$

$$(1.2) \quad \sum_{k=0}^{\infty} a_{nk} = 1 \quad (n = 0, 1, \dots).$$

és

$$(1.3) \quad \lim_{n \rightarrow +\infty} a_{nk} = 0 \quad (k = 0, 1, \dots),$$

legyen s_n a szummálandó sorozat és

$$(1.4) \quad t_n = \sum_{k=0}^{\infty} a_{nk} s_k$$

a transzformált sorozat, amelynek konvergenciája esetén az s_n sorozatot az A mátrix által definiált szummációs eljárással szummálhatónak nevezzük. A $\{t_n\}$ sorozat elemei a következőképpen interpretálhatók: legyen v_n egy nemnegatív egész értékű valószínűségi változó, amely a k értéket a_{nk} valószínűséggel veszi fel, mely esetben az (1.3) feltétel azt fejezi ki, hogy v_n sztochasztikusan $+\infty$ -hez tart, ha $n \rightarrow +\infty$. t_n felfogható, mint s_{v_n} várható értéke. Ez az interpretáció különösen a HAUSDORFF-féle és HENRIKSSON-féle szummációs eljárások esetében hasznos, és lehetővé teszi ismert összefüggések egyszerű bizonyítását oly módon, hogy azokat jólismert binomális, ill. Poisson-eloszlások keverésére vonatkozó tételekre vezethetjük vissza. A 4. §. egy divergens sorozat határeloszlásával foglalkozik, és az ennek segítségével értelmezhető szummációs eljárásokkal. Az 5. §. a szerző egy sor összes véges rész-összegeinek sorozataira vonatkozó sejtésével foglalkozik, amely a következő alakra hozható: ha

$$\lim_{x \rightarrow 1-0} \sum a_k \frac{x^{2k}}{1+x^{2k}} = s$$

létezik, akkor a $\sum a_k$ sor konvergens, és összege $2s$. Ezt az állítást a szerzőnek ezideig csak a $\lim a_k = 0$ kiegészítő feltevés mellett sikerült bebizonyítania.

МЕТОДЫ СУММИРОВАНИЯ И ИХ ТЕОРЕТИКО-ВЕРОЯТНОСТНАЯ ИНТЕРПРЕТАЦИЯ

A. RÉNYI

Резюме

Пусть $A = (a_{nk})$ бесконечная матрица, для элементов которой имеют место следующие соотношения

$$(1.1) \quad a_{nk} \geq 0 \quad (n, k = 0, 1, \dots),$$

$$(1.2) \quad \sum_{k=0}^{\infty} a_{nk} = 1 \quad (n = 0, 1, \dots),$$

$$(1.3) \quad \lim_{n \rightarrow +\infty} a_{nk} = 0.$$

Пусть s_n — последовательность числа, которую хотим суммировать и посмотрим последовательность

$$(1.4) \quad t_n = \sum a_{nk} s_k.$$

Если t_n сходится к пределу s , то мы скажем, что последовательность s_n суммируемый к s с методом суммирования определённым матрицей \mathbf{A} .

Эту методу можно истолковать на языке теории вероятностей следующим образом: пусть v_n неотрицательная целочисленная случайная величина которая принимает значение k с вероятностей a_{nk} . Тогда t_n равно математическому ожиданию от s_{v_n} . Это истолкование особенно полезно при рассмотрении методов суммирования типа HAUSDORFF и HENRIKSSON и делает возможным сводить доказательство нескольких теорем об этих методах суммирования на хорошо известных элементарных фактов относительно смеси биномиальных и Пуассонских распределении вероятностей. § 4 занимается методами суммирования определенных с помощью предельных распределении расходящихся последовательностей.

В § 5 формулируется недоказанная гипотеза о том что если имеет место

$$(5.5) \quad \lim_{x \rightarrow 1-0} \sum_{k=0}^{\infty} a_k \frac{x^{2^k}}{1+x^{2^k}} = s,$$

то ряд $\sum_{k=0}^{\infty} a_k$ сходится и имеет сумму $2s$; доказывається, что это верно если кроме (5.5) предполагается что a_k стремится к нулю.

A kiadásért felel az Akadémiai Kiadó igazgatója

Műszaki szerkesztő: Pataki Ferenc

A kézirat beérkezett: 1959. XI. 10. Terjedelem: 17,50 (A/5) ív

50.431/60 — Akadémiai Nyomda, Budapest — Felelős vezető: Bernát György

СОДЕРЖАНИЕ

FINNEY, D. J. : О рентабельности экспериментирования	203
Линник, Ю. В. : Пять лекций о некоторых вопросах теории чисел и теории вероятностей	225
GRÄTZER, Gy.—SCHMIDT, E. T. : Одна теорема об ассоциативности для альтернативных колец	259
KREKÓ, B. : Решение систем линейных уравнений методом симплекса	265
RÓZSA, P.—SÁRKÁNY, Gy. : Определение числа теоретических ступеней противоточных разделяющих процессов химической технологии, II.	277
BOSZNAJ, Á. : Один метод исследования возбужденных колебаний колеблющихся систем с несколькими степенями свободы и несколько его приложений	299
CSÁKI, E. : О двух видоизменениях статистики Wilcoxon-a	313
HASZPRA, O. : Несколько замечаний в связи с погрешностями номограмм	321
ADLER, Gy. : Охлаждение зернистых веществ встречным потоком	327
ADLER, Gy.—FREUD, G. : Применение операционного исчисления от Mikusiński к решению одного уравнения в частных производных	367
KOSIK, P.—SALLAY, M.—ZIMÁNYI, M. : Проблемы теплопроводности в случае сложных граничных условий	377
RÉNYI, A. : О связных графах, I.	385
RÉNYI, A. : Методы суммирования и их теоретико-вероятностная интерпретация	389

INDEX

FINNEY, D. J.: The economic efficiency of experimentation.....	203
LINNIK, JU. V.: Five lectures on some topics in number theory and probability theory	225
GRÄTZER, GY.—SCHMIDT, E. T.: An associativity theorem for alternative rings....	259
KREKÓ, B.: Solution of a system of linear equations with the simplex method....	265
RÓZSA, P.—SÁRKÁNY, GY.: Determination of the theoretical stage-number of coun- tercurrent distributive chemical unit operations, II.	277
BOSZNAY, Á.: Eine Methode zur Untersuchung der erzwungenen Schwingungen von Schwingungssystemen mit mehreren Freiheitsgraden und deren Anwen- dungen	299
CSÁKI, E.: On two modifications of the Wilcoxon-test	313
HASZPRA, O.: Einige Bemerkungen zur Berechnung der Nomogramablesefehler	321
ADLER, G.: Réfrigération de retour des matières granuleuses	327
ADLER, G.—FREUD, G.: Una applicazione del calcolo degli operatori di Mikusiński per la risoluzione d'una equazione alle derivate parziali	367
KOSIK, P.—SALLAY, M.—ZIMÁNYI, M.: Problèmes de la propagation de la chaleur avec des conditions aux limites composées	377
RÉNYI, A.: On connected graphs, I.	385
RÉNYI, A.: Summation methods and probability theory	389